CITY UNIVERSITY OF HONG KONG
香港城市大學

Predictive Processing of Multiple Acoustic Features in the Auditory Cortex
聽覺皮層中多種聲學特徵的預測處理

Submitted to
Department of Neurosciences
神經科學系

in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
哲學博士學位

by

An Hyunjung

June 2021
二零二一年六月

Abstract

Brain responses that differentiate the processing of unexpected deviant stimuli from expected standard stimuli are often quantified as mismatch negativity (MMN). The MMN (a difference in event-related potentials based on non-invasive recordings evoked by sensory deviants and standards) has long been recognized to reflect novel stimulus processing. Previous studies in humans (Phillips et al., 2015) suggest that MMN is related to deviance detection based on sensory prediction violations. However, it has also been suggested that predictions of stimulus contents ("what") vs. stimulus timing ("when") have different putative mechanisms (Auksztulewicz et al., 2018b). While the latter study raises the possibility that violating predictions of different stimulus attributes can be dissociable, it is unknown whether this functional specialization extends to different content-based predictions, such as acoustic pitch, location, duration, and spectral composition. Furthermore, since deviant detection rests on extracting a memory representation of recent stimulus statistics, the MMN can be used as an indirect measurement of sensory memory formation. However, other more direct measurements, such as intracranial recordings from animal models, are needed to elucidate the neural mechanisms of auditory memory formation.

In this thesis, I investigated whether violations of acoustic predictions based on different stimulus features (pitch, duration, location, or formant) are reflected in differences in neural activity in human and animal models. The first two experiments used a roving oddball experimental paradigm, adapted from a previous study (Garrido et al., 2008). In the first experiment, I investigated whether feature-specific differences in the spatial distribution of MMN responses can be mapped onto different cortical regions using electrocorticography

(ECoG) recording in the rat auditory cortex. We demonstrated that MMN could be observed following the violation of four independent acoustic features, but that these MMN responses show a large degree of heterogeneity not only across different acoustic features but also across individual animals.

In the second experiment, I applied the same paradigm to humans using electroencephalography (EEG). As in the rat study, we found significant mismatch responses following prediction violations of four independent acoustic features. While no significant differences were observed between MMN signals corresponding to different stimulus features in a traditional univariate analysis, acoustic feature information could be decoded from the fine-grained topography of mismatch responses based on multivariate analysis. Consistent with previous studies, the results indicated that deviant detection along different stimulus features could be linked to differences in the spatial distribution of neural responses.

In the third experiment, I used a direct manipulation of auditory memory formation and recorded auditory cortical activity in anesthetized rats to infer the neural correlates of auditory pattern learning, differentiating between repeated sequences and novel sequences. While no significant differences between re-occurring and new sequences were observed in event-related potentials (ERPs), when comparing responses in the time-frequency domain, we found robust differences between re-occurring and new sequences. This result was most prominent in the beta frequency band, indicating a neural correlate of learning formation in passively listening and anesthetized animal models.

In conclusion, the thesis shows that the neural correlates of sensory deviance detection depend on the type of sensory feature whose violation underlies deviant stimulus

presentation and that the auditory cortex is susceptible to auditory memory formation both through indirect measurements (mismatch responses) and more direct measurements (repeated presentation of the same stimulus sequence).

CITY UNIVERSITY OF HONG KONG

Qualifying Panel and Examination Panel

Surname	An
First Name	HyunJung
Degree	Ph.D.
College/Department	Department of Neuroscience

The Qualifying Panel of the above student is composed of
Qualifying Panel members(s)

Supervisor:	Department of Neuroscience
Prof. Jan W.H.Schnupp	City University of Hong Kong
Dr. Sungchil Yang	Department of Neuroscience
	City University of Hong Kong
Dr. Geoffrey C. Y. LAU	Department of Neuroscience
	City University of Hong Kong

This thesis has been examined and approved by the following examiners

Prof. Manuel S. Malmiera	Department of Cell biology and pathology
	Medical School University of Salamanca
Dr. Chan L H Leanne	Department of Electrical Engineering
	City University of Hong Kong
Dr. Tse Chun Yu	Department of Social and Behavioural Science
	City University of Hong Kong

DECLARATION

I, HyunJung An, declare that this thesis entitled, ”Predictive processing of multiple acoustic features in the auditory cortex” represents my original work and the contents of this thesis has never been submitted to this University and other institutions for a degree or any other qualifications in the form of thesis or other report.

HyunJung An

DEDICATION

To my family

(An kyungman, Moon kyunghee, An Hyunjune, An sanghyeon, Jung seunghun)

And

My grandparents

(Kim sookyoon, Moon Jongyel)

To all of them, I dedicate this work

ACKNOWLEDGMENTS

I would like to thank all the people who have made this thesis possible.

Jan W.H. Schnupp, of course, I want to thank you in particular. Without your supervision and your support, none of this would have been possible. I recall that already the first interview for the Ph.D. position felt more like a pleasant discussion of neuroscience among colleagues than a job interview. To my surprise, you offered me the position still on the same day. That moment is the luckiest day in my life. I have since come to know you as an excellent supervisor who is both knowledgeable and supportive. I could not have asked for a better supervisor. Thank you for everything!

Ryszard Auksztulewicz, I cannot thank you enough for everything. Sometimes you were guiding me on finding my way (figuratively and literally) at the auditory neuroscience. You were not only my office-mate to discuss science but also my mentor. You made the office into a second home, and I am happy to share our time during my Ph.D. HiJee, you, as much as Ryszard, made the office into my second home to share the idea and all kinds of gossip. I recall many late evenings in the office. I have enjoyed working with you immensely, both on your and my projects; I hope that more collaborations will come in the future. Of course, Fei made the office the fun place it is. I am always happy to hear your insightful thoughts, which often helped me to consider another perspective. And I was always enjoying our chats about work and all sorts of same concerns.

Of course, I also want to thank all the auditory neuroscience lab members. Alex, Ambika, Bruno, Chloe, Cecilia, Drew, Jacinta, and Shiyi. Thank you all for making the lab an inspiring and hospitable place.

Thank you to my family for supporting me throughout all the years, especially my mom and father, to give me endless trust and provide amazing support for my choices. I learned patience from my dad and challenges from my mom. I am so proud that you are my parents. And, thank my grandmother, who is the strongest and most beautiful woman in the world. Without her, I probably would not have finished this long journey. I will become a proud granddaughter for loving and supporting me unconditionally. I would also like to express my deep gratitude to my sister. Sometimes you were like a friend and a counselor in life. Thank you for stay with me whenever I have troubles. Thanks to my little brother. As a colleague in the same major (audiology), I always felt like you are still a young boy, but I am proud of you for having your business. I will always try to be a big supporter.

Lastly, my grandfather passed away during my Ph.D. It's so sad that we couldn't be together on the day that the whole process was over. But I believe you have always protected and supported me from heaven. I love you so much. I miss you.

List of Abbreviations

A1	Primary auditory cortex
AAF	Anterior auditory field
ABRs	Auditory brainstem responses
AC	Auditory cortex
AEPs	Auditory evoked potentials
CSI	Common SSA index
CV	Consonant-vowel
DCN	Dorsal cochlear nuclei
DRCs	Dynamic random chords
ECoG	Electrocorticography
EEG	Electroencephalography
ERPs	Event-related potentials
F0	Fundamental frequency
fMRI	Functional magnetic resonance imaging
FRA	Frequency response area
FSL	Frist spike latency
GLM	General linear model
IC	Inferior colliculus
IFG	Infeior frontal gyrus
ILD	Interaural level difference
ISIs	Inter-stimulus intervals

ITD	Interaural time difference
ITPC	Inter-trial phase coherence
LEP	Local field potential
LLN	Lateral lemniscal nuclei
LSO	Lateral superior olive
MEG	magnetoencephalography
MEPs	Memory-evoked potentials
MGB	Medial geniculate body
MGBd	Medial geniculate body dorsal
MGBm	Medial geniculate body medial
MGBv	Medial geniculate body ventral
MMN	Mismatch negativity
MNTB	Medial nucleus of trapezoid body
MSO	Medial superior olive
PAF	Posterior auditory field
PCA	Principal component analysis
PFC	Prefrontal cortex
PSTH	Peri-stimulus time histogram
REM	Rapid eye movement
RP	Repetition positivity
RS	Repeated sequence
SRAF	Suprahinal auditory field
SSA	Stimulus-specific adaptation

STG	Superior temporal gyrus
VAF	Ventral auditory field
VCN	Ventral cochlear nuclei
WN	White noise

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Importance of predictive processing in auditory cortex

The ability to automatically detect signal changes is essential to survival for any organism. For example, an animal living in the forest can distinguish a relatively constant sound background, such as leaves blowing in the wind, from a sudden novel sound, such as a cracking branch, which might indicate the presence of a predator nearby. Distinct brain responses to novel sounds (as opposed to expected sounds) are known as mismatch negativity (MMN) in humans and are thought to reflect this ability of auditory change detection. MMN was observed at first by Näätänen et al. (1978) in the human electroencephalogram (EEG) as a late negative deflection in auditory evoked potentials (AEPs) following a presentation of deviant stimuli. They found that the neural response was attenuated by the sound sequence of repetitive stimulation while a presentation of an unexpected stimulus enhanced the neural response. The MMN amplitude is typically observed at approximately 100 - 250 ms post-stimulus in humans (Javitt et al., 1994, 2008; Nattanen et al., 1978, 2007; Näätänen and Alho, 1995b; Näätänen and Michie, 1979)

Two main hypotheses have been proposed concerning the neural mechanisms of the MMN: the model adjustment hypothesis and the neural adaptation hypothesis. According to the model adjustment hypothesis (Näätänen and Alho, 1995; Näätänen et al., 2005), incoming sounds (deviants) are compared with the neural memory trace established by prior exposure of sounds (standards), which indexes the sound change detection of a mismatch between properties of deviant and standard sounds. In contrast, the neural adaptation hypothesis assumes that MMN does not reflect a higher level comparison process but rather results from neural adaptation to

repeated sounds (standard). According to the latter view, repeated sound presentation attenuates neural responses in the auditory cortex due to synaptic depression and inhibition (May et al., 1999; May and Tiitinen, 2010). However, it can be argued that both of these models are too simple to explain the ability to predict more complex changes in the acoustic environment.

Alternatively, predictive coding (Friston & Kiebel, 2009; Rao & Ballard, 1999), one of the most widely accepted integrative brain function theories, assumes that the brain continuously updates the internal model to minimize the error between the incoming input and the model. In this view, the brain is considered a hierarchically organized system where each level adjusts between the bottom-up sensory inputs and top-down predictions (Garrido et al., 2009). In this framework, MMN is typically interpreted as reflecting the discrepancy between top-down predictions and bottom-up inputs.

While the predictive coding interpretation of the MMN is relatively recent, MMN was also subject to extensive neuroscientific research prior to the formulation of the predictive coding framework. Previous studies have investigated MMN properties across a wide range of stimulus conditions, including deviant stimulus type (i.e., violating predictions of different stimulus features such as pitch, duration, location, identity, etc.), deviant probability, and interstimulus interval (Korzyukov et al., 1999; May et al., 1999; Korzyukov et al., 2003; Jaaskelainen et al., 2004). However, non-invasive recordings in humans using EEG techniques have a limited spatial resolution, making it difficult to answer more detailed questions about mismatch detection mechanisms along the auditory pathway.

Therefore, many researchers have turned to animal models, such as rodents, which offer the possibility of direct recordings from neuronal populations and single neurons. In the rodent models,

the MMN amplitude is typically observed at approximately 50 – 150 ms post-stimulus (Ruusuvirta et al., 1998; Lazar and Metherate, 2003; Astikainen et al., 2006; Roger et al., 2009a; Astikainen et al., 2011; Nakamura et al., 2011). The MMN latency difference between humans and animals is attributed to the brain's size and complexity (Javitt et al., 1992). Despite this difference, mismatch response has been shown to occur following manipulations of multiple acoustic features such as syllable patterns (Ahmed et al., 2011; Mahmoudzadeh et al., 2017), duration (Lipponen et al., 2019), and frequency (Eriksson and Villa, 2005). While these studies in rodents suggest that violations of different acoustic features can be used to elicit mismatch responses, recent studies in humans suggest that different neural populations and mechanisms mediate prediction and prediction error corresponding to different stimulus features (Auksztulewicz et al., 2018b; Stefanics et al., 2019). However, since the previous studies in rodent models did not address differences in mismatch responses along different acoustic features, a critical question remains whether mismatch signaling is a domain-general or domain-specific process of detecting deviant stimuli.

Domain-general learning theories posit that learning different types of new information may be processed in the same way through the same brain areas and that learning in different domains may function interdependently (Sloutsky, 2010; Li et al., 2014). Conversely, domain-specific learning theories posit that humans detect different information types differently through specialized brain networks and mechanisms (Cosmides et al., 2010; Cosmides and Tooby, 2013). Functional imaging studies have shown that simple stimuli such as pure tones are sufficient to drive activity in the auditory core (Heschl's gyrus), whereas more complex stimuli such as bandpass noise or speech responses drive activity in higher-order (belt and para-belt) regions (Binder et al., 2000; Scott et al., 2000; Wessinger et al., 2001; Hickok and Poeppel, 2007). These

results support a hierarchical auditory processing model in the human auditory cortex with different regions playing distinct functional roles. In principle, the hierarchical organization of the auditory cortex has been linked to distinct regional distribution and independent processing of different stimulus features. Several studies (Schröger, 1995; Paavilainen et al., 2001; Phillips et al., 2015; Rosburg et al., 2018) support the evidence that separable neural mechanisms mediate mismatch responses to violations of different stimulus dimensions. However, they did not consider that physical differences between stimuli along acoustic features should be individually adjusted to yield similar behavioral responses. Therefore, Chapter 2 in this thesis will address the regional specificity of neural mismatch responses to violations of different content-based predictions in the rat model. Furthermore, testing the evolutionary conservation of domain-specificity of MMN signalling, Chapter 3 will answer the question whether MMN responses in humans differ depending on whether sudden stimulus change occurs in pitch, duration, location, or vowel identity using behaviorally adjusted deviant sounds.

Finally, to understand the interrelations between MMN and sensory memory, it is essential to investigate the underlying neural mechanisms in the auditory cortex. MMN is elicited by a mismatch between the deviant auditory input and sensory-memory trace representing the standard stimuli. Earlier studies support the notion that MMN increases with the number of standard stimulus repetitions (Näätänen, 1992; Imada et al., 1993). This result suggested that MMN is a neurophysiological index for the echoic memory trace. However, MMN could not directly explain echoic trace formation because MMN depends on preceding standard formation. Previous electrophysiological studies have investigated how neurons adapt to re-occurring sounds to understand memory and adaptation processes by using a simplified experimental paradigm, in which the presentation of standard sounds (pure tones) is disrupted by a presentation of a deviant

sound (Garrido et al., 2009a; Malmierca et al., 2014; Nieto-Diego and Malmierca, 2016). However, realistic auditory scenes typically do not just include single sounds, but occasionally repeated appearances of more complex stimuli. As a result, studies using single sounds have limited ecological validity when explaining auditory pattern learning. Therefore, Chapter 4 in this thesis will focus on the physiological correlates of auditory pattern learning using repetitive exposure to a specific complex sound sequence.

1.2 Neural pathways in the auditory system

Auditory neurons encode and decipher the spectral, spatial, and temporal properties of sound. This chapter describes the auditory pathway that extends from the external ear, via the cochlea and several synaptic relays, to the cerebral cortex. The hearing process relies on three sequential stages along the pathway: the outer ear, middle ear, and inner ear. The outer and middle ear collect sound and physically travel the sound toward the inner ear that consists of the cochlea. The cochlea has three fluid-filled sections (i.e., the scala media, scala tympani, and scala vestibuli). Corti's organ is located in this duct on the basilar membrane and transforms mechanical waves into neurons' electric signals. The chemical difference between the fluids endolymph and perilymph fluids is vital for the inner ear's function due to electrical potential differences between potassium and calcium ions. Therefore, the auditory system transforms a wide range of mechanical signals into a complex electrical signal series in the central nervous system (Figure 1).

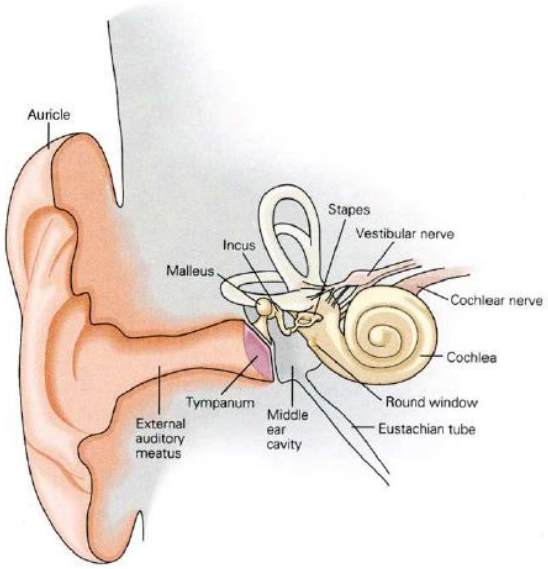

Figure 1. The structure of the human ear.
There are three major parts of the overall ear structure. The external ear, or auricle, collects sounds into the external auditory meatus. The sound vibrates the tympanum. These vibrations are conveyed to the middle ear bones; the malleus, the incus, and the stapes. The vibration of stapes is transmitted to the cochlea in the inner ear (Schultz, 2001).

The neurons of the cochlear nuclei in the brainstem are the first central processors of auditory information, providing inputs to all the major brainstem and midbrain auditory nuclei. The synaptic connectivity pattern of neural networks in each part of the cochlear nuclear complex is an essential determinant of their information processing role (Figure 2).

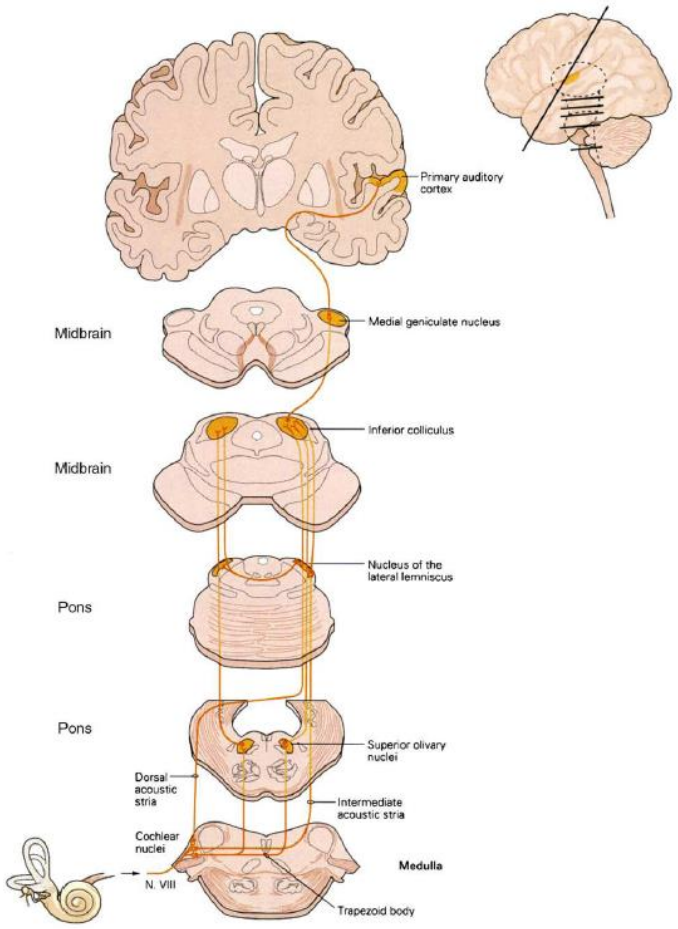

Figure 2. The central auditory pathway

The figure outlines the central auditory pathway from the cochlear nucleus to the auditory cortex. Neurons in the cochlear nucleus send their parallel axons along three main pathways; the dorsal acoustic stria, the intermediate acoustic stria, and the trapezoid body. The first construction of binaural representations is in the superior olivary nucleus, which is involved in sound localization—post-synaptic axons from the superior olivary nucleus project to the lemniscus. The inferior colliculus converges multisensory information. The geniculate axons terminate in the auditory cortex, providing an interface between sensory and cognitive networks (Schultz, 2001).

Cochlear Nuclei

The first relay of the primary auditory pathway occurs in the cochlear nuclei in the brain stem, which receives spiral ganglion axons (auditory nerve). The major role of these nuclei is to preserve

the order of input according to the topographic organization of frequency encoding, established in the cochlea and specialized post-synaptic neurons that process the auditory nerve's output (Osen, 1969; Cant and Benson, 2003). This region is anatomically split into two regions in mammals, the dorsal cochlear nuclei (DCN) and ventral cochlear nuclei (VCN). From the DCN, most fibers cross the midline and ascend in the contralateral lateral lemniscus. Other fibers ascend in the ipsilateral lateral lemniscus (Brawer et al., 1974). From the VCN, some fibers also ascend in the lateral lemniscus bilaterally. The second primary relay in the brain stem is in the superior olivary complex: most of the auditory fibers synapse are there, having already crossed the midline (Lee et al., 1996).

Superior Olivary Complex

The olivary complex consists of three primary nuclei – the medial superior olive (MSO), lateral superior olive (LSO), and medial nucleus of trapezoid body (MNTB) – and several smaller preolivary nuclei (Guinan and Li, 1990; Yin and Chan, 1990; Kavanagh and Kelly, 1992). They have several complementary physiologic roles. Neurons in the MSO encode phase relationships and delay sensitivity from the two ears, mainly from lower frequencies; these signals are essential for accurate spatial localization (Yin and Chan, 1990). Neurons in the LSO converge monaural input of the cochlear nuclei to derive binaural signals sensitive to intensity difference and send this information from the lateral lemniscal nuclei (LLN) to the inferior colliculus (IC) (Kavanagh and Kelly, 1992). Finally, MNTB contributes to creating binaural subtypes via their inhibitory input to the LSO (Guinan and Li, 1990). The construction and variety of binaural interactions allow for emergent functions to directly relate to a particular neural circuit (Covey et al., 1991).

Inferior Colliculus

The inferior colliculus (IC) is a dome-shaped structure, among the largest auditory nuclei in the vertebrate brain. It is virtually an obligatory synaptic terminus for ascending input to the medial geniculate body (MGB) (Aitkin and Reynolds, 1975). Besides, the IC's connectivity pattern suggests its critical roles in both the auditory ascending and descending pathway. The IC contains three principal divisions: central nucleus, lateral nucleus, and dorsal cortex. Each has several nuclei that differ in neuronal structure (Morest and Oliver, 1984), connectivity (Rockel and Jones, 1973) and functional role (Aitkin et al., 1970). The central nucleus is exclusively auditory (Aitkin et al., 1994) and is essential for normal hearing (Jenkins and Masterton, 1982). The lateral nucleus is multisensory and the target of considerable non-auditory input (Morest and Oliver, 1984). Lastly, the dorsal cortex receives most of its projections from the cerebral cortex (Winer et al., 1995).

The Medial Geniculate Body

The primary sensory nucleus of the medial geniculate body (MGB) possesses three subdivisions: ventral (MGBv), dorsal (MGBd), and medial (MGBm) bodies (Clerici and Coleman, 1990; Winer et al., 1999; Malmierca, 2003). An earlier study (Graybiel, 1973) defined two sensory conduction routes referred to as lemniscal and lemniscal-adjunct (or non-lemniscal) systems. Since then, auditory research is widely used to easily classify and understand the role of multiple subdivisions (Hu, 2003; Jones, 2003; Lee and Winer, 2008). The lemniscal pathway consists of the central nucleus IC, the MGBv, and the primary auditory cortex (AC) in the rat study (Parras et al., 2017). The lemniscal division is thought to be in charge of accurately relaying sensory input and tends to be sharply tuned to frequency response. In addition, lemniscal neurons showed a consistent response to the sound including shorter latencies and higher spontaneous activity (Malmierca et

al., 2015). The other system referred to as the non-lemniscal pathway includes the rostral and dorsal part of IC, MGBd-MGBm, and secondary area of AC in rodents (Parras et al., 2017). This pathway consists of a belt of broadly tuned neurons that receive input from lemniscal and from other non-lemniscal stations. This connectivity reflects their comparatively longer latencies and breadth of frequency response (Saldana et al., 1996; Malmierca and Ryugo, 2011). Therefore, the non-lemniscal pathway seems to form a higher-order pathway processing, handling more complex aspects of the auditory scene and being required to account for the emergence of deviance-detection activity (Carbajal and Malmierca, 2018).

Primary Auditory Cortex

The auditory cortex plays a critical role in sound processing as it comprises multiple receptive fields with a hierarchical organization. Three major subdivisions are identified in primates referred to as the core, belt, and para-belt regions. The cells in the core areas have relatively short response latencies and well-defined frequency responses. In the caudal area, neurons are best activated by high-frequency tones whereas rostral area neurons are activated by low-frequency tones (Merzenich and Brugge, 1973; Imig et al., 1977; Aitkin et al., 1986; Morel and Kaas, 1992; Morel et al., 1993; Kosaki et al., 1997). The auditory belt surrounds the core with dense interconnections with the core. The belt appears to receive only sparse inputs from the ventral part of the medial geniculate complex with most of its thalamic inputs coming from the dorsal and medial division of the complex (Rauschecker et al., 1997). The para-belt is interconnected with adjunct portions of the temporal and parietal lobe and several frontal lobe regions (Yeterian and Pandya, 1989; Kosmal et al., 1997; Romanski et al., 1997). Para-belt neurons are unlikely to respond to pure tones and prefer complex stimuli such as vocalizations. The previous functional imaging studies of the

human auditory cortex have suggested the hierarchical organization in which the simple stimulus such as pure tone is sufficient to drive activity in the auditory core whereas complex stimuli such as bandpass noise or speech produced activity in surrounding auditory cortical regions (Binder et al., 2000; Scott et al., 2000; Wessinger et al., 2001; Hickok and Poeppel, 2007).

Moreover, deviance detection and prediction error are widespread in the auditory cortex. The primary auditory cortex is the first station in the lemniscal pathway to exhibit deviance detection reliably. The predictive error seems to account for about 25% of the overall deviance-detection activity recorded in the lemniscal AC, which is a more significant proportion than in non-lemniscal divisions of the subcortical area (Parras et al., 2017). Therefore, in this thesis, we measured the auditory cortex differences between primary and higher-order regions map.

1.3 Mismatch negativity (MMN) and stimulus-specific adaptation (SSA)

Mismatch negativity (MMN)

When presented with novel input, the brain responses reveal a unique pattern. Neural activity waveforms typically show suppression (attenuation) of responses to the familiar or expected stimuli whereas an enhanced response reflects novel or unexpected stimuli. This response pattern, when observed in local field potential (LFP) or electroencephalography (EEG) signals, is called the mismatch negativity (MMN) and comprises an auditory event-related potential (ERP) calculated by subtracting the response to a standard or repeated stimulus from the responses to a deviant or novel stimulus (see Figure 3 (a)). MMN is interpreted as a brain response reflecting automatic deviance detection. In traditional MMN-inducing paradigms, such as the oddball paradigm, the probability of standard stimulus presentation is, by definition, much higher (ca. 90%)

than the probability of deviant stimulus presentation (ca. 10%) (see Figure 3 (b)). The peak latency is approximately 100 - 250 ms (in the N1 range and later) relative to the onset of the deviant stimulus. Interestingly, MMN responses might differ depending on the type of stimuli. Previous research studied the modulation of MMN amplitude by the feature of stimulus deviance such as pitch, duration, formant, and location (Alho et al., 1998; Grimm et al., 2004; Ahmed et al., 2011; Komatsu et al., 2015). The details of multi-feature deviance detection are described in section 1.5. According to source reconstruction studies, the MMN neural generators are typically localized within the auditory cortex, including the primary auditory cortex (A1) and superior temporal gyrus (Hari et al., 1984; Giard et al., 1990; Näätänen and Alho, 1995; Alho et al., 1998). However, the correlation of neural mechanisms and specific brain regions remains unresolved.

Recently, MMN research mainly considers the underlying neural mechanisms (Garrido et al., 2009b). There are two main hypotheses proposed to explain it. The first hypothesis has been referred to as the "model adjustment" hypothesis. The second has been referred to as the "neural adaptation" hypothesis (Näätänen et al., 2005; May and Tiitinen, 2010). These hypotheses are discussed in detail in section 1.4.

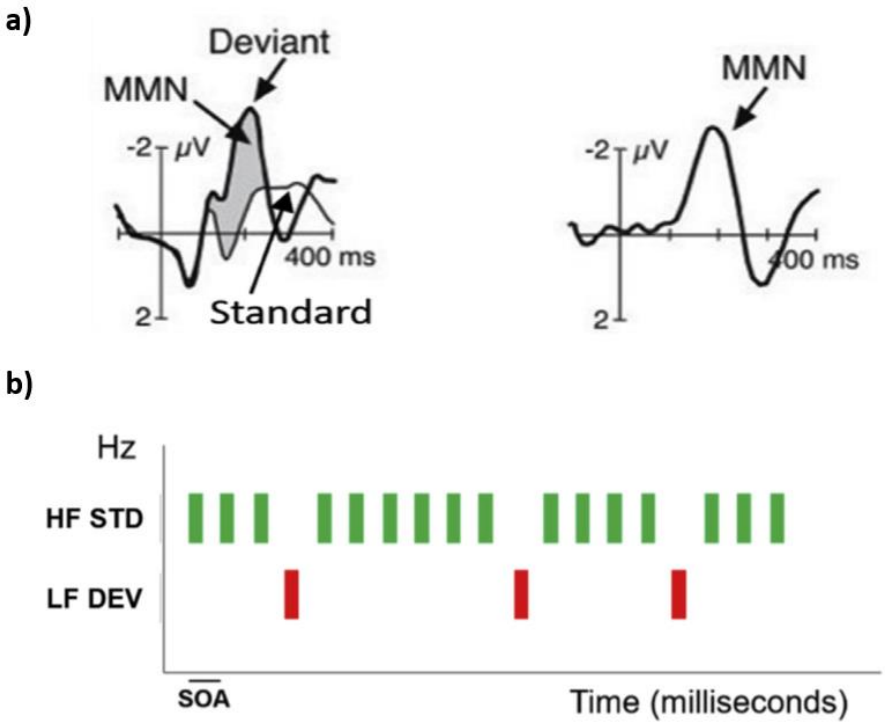

Figure 3. Schematic representation of the mismatch negativity (MMN) and oddball paradigm

a) MMN is marked as a shaded area, which is calculated as a standard response (thin line) subtracted from the deviant response (bold line). b) Descending oddball paradigm where expected high-frequency stimuli serve as standard sounds (green) and unexpected low-frequency stimuli serve as deviant sounds (red) (Kujala and Näätänen, 2001).

Stimulus-specific adaptation (SSA)

Using the same oddball sequence as in MMN studies, a similar deviance-sensitive response has been identified in the response of single neurons along the auditory pathway in animal models. These neurons show progressively reduced response to repetitive sounds, which is restored when exposed to novel sounds. This single-neuron phenomenon is referred to as stimulus-specific adaptation (SSA) (Movshon and Lennie, 1979). SSA is quantified as the index of change in a

neuron's response firing rate between deviant and standard stimuli. This index is referred to as the common SSA index (CSI), which is calculated as

$$CSI = [d(f1)+d(f2) - s(f1) - s(f2)]/[d(f1) + d(f2)+s(f1)+s(f2)]$$

Here, d(f) and s(f) denote spike rate to two frequencies (*f1* and *f2*) for deviant (d) or standard (s) stimulus. CSI measures the proportional difference in spike rate between deviant and standard. When the deviant stimulus-response is stronger, the CSI value is positive (see Figure 4(a)).

Compared to MMN, SSA is typically observed at earlier latencies – for instance, one previous SSA study (Harms, 2016) showed the difference in spiking responses peaking at 50-100 ms. SSA is observed subcortically in the IC (Perez-Gonzalez et al., 2005; Reches and Gutfreund, 2008; Malmierca et al., 2009; Lumani and Zhang, 2010; Netser et al., 2011; Zhao et al., 2011), and the MGB (Anderson et al., 2009; Yu et al., 2009; Antunes et al., 2010) as well as in the primary AC (Ulanovsky et al., 2003; Ulanovsky et al., 2004a; von der Behrens et al., 2009; Taaseh et al., 2011). Among those above, SSA to frequency deviance is stronger in the non-lemniscal regions (IC) such as in the dorsal, rostral and lateral cortices than in the central nucleus (Perez-Gonzalez et al., 2005; Malmierca et al., 2009; Lumani and Zhang, 2010; Malmierca et al., 2011; Duque et al., 2012) (Figure 4(b)). Previous studies have observed sensitivity to intensity and duration deviance in the AC, but it is not as robust as responses to frequency deviance (Ulanovsky et al., 2003; Farley et al., 2010).

Another finding is that SSA is not homogeneously distributed within the frequency response in IC neurons. In a study of the IC in the rat (Duque et al., 2012), the authors compared SSA levels at multiple combinations of frequency and intensity using single-unit recordings. They found that adapting neurons show stronger SSA at high frequency and low sound intensity. According to a

line of previous studies (Malmierca et al., 1993; Malmierca et al., 1995; Malmierca et al., 2011), the morphological differences reflected differences in IC location. They found that dorsal and rostral IC regions possess widespread dendritic arbors and broader frequency tuning than the central IC region (Duque et al., 2012). Besides the encoding deviance by spike count, IC neurons could also encode deviance information through spike timing. The first spike latency (FSL) evoked by the deviant stimulus is shorter than the standard stimulus (Malmierca et al., 2009; Zhao et al., 2011; Duque et al., 2012). This is referred to as the latency adaptation, which seems to be a uniquely subcortical phenomenon (Malmierca et al., 2009; Antunes et al., 2010; Duque et al., 2012). Thus, these results indicate that temporal coding seems to play a vital role in the signaling of deviance.

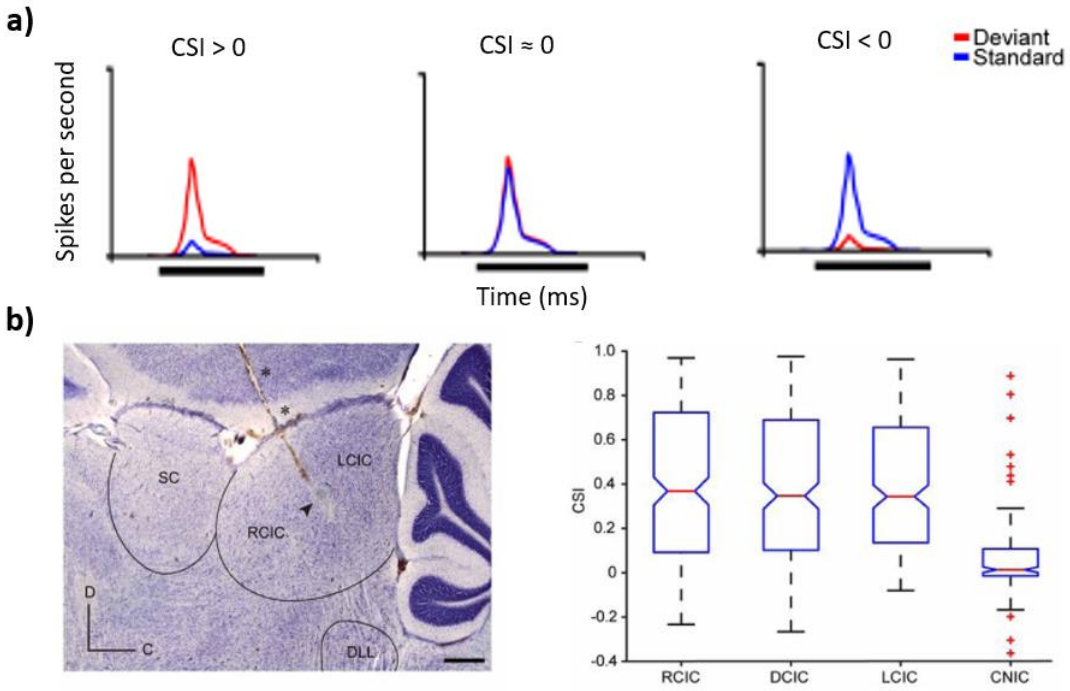

Figure 4. Illustration of Stimulus-Specific Adaptation (SSA)

a) The common SSA index (CSI) is used to quantify the SSA neuronal responses. Three peri-stimulus time histograms (PSTHs) represent the response of the deviant (red line) and standard (blue line) for three different CSI values. b) Photomicrography showed the anatomical location of SSA in the IC. The plot represents a sagittal section of the IC with electrode track (asterisks) and electrolytic lesion (arrowhead). C (caudal); D (dorsal). The right panel (box plot) shows that CSI value depends on anatomic regions. The blue box represents the 25^{th} - 75^{th} percentile, and the red crosses are outliers (Duque et al., 2012).

1.4 Different ways of understanding deviance detection

1.4.1 Model adjustment hypothesis

While the MMN has been subject to intensive research regarding its spatial distribution, latency, and sensitivity to stimulus characteristics, the hypothesized mechanisms have been debated. Earlier studies (Winkler and Czigler, 1998) suggested the model adjustment hypothesis to explain the mismatch between the standard stimuli and deviant stimuli. According to this theory, the MMN is thought to reflect an online update of the prior perceptual model established based on the previous experience and compared with the actual auditory input. As a result of the update, the neural response to the deviant increases relative to the neural response to standard stimuli (Figure 5.1). In agreement with this theory, a previous study demonstrated deviant detection updated model of the sensory input in auditory processing. (Sussman and Winkler, 2001). In this study, the researchers examined the effect of the single deviants on the response to double deviant in a contextual change paradigm. The deviants were presented in two ways; one single tone (single deviant) and two successive tones (double deviant) had a higher pitch than standard. The responses of both the first and the second deviant evoked significant MMN, but the MMN evoked by the first differed from that evoked by the second. The first deviant of the double deviant showed MMN both at the beginning and end of the segment, but the second deviant of a double deviant elicited

MMN only at the end of the segment. This result indicated that the auditory system attended double deviant at the beginning (integration) and only after sufficient information is obtained can different stream emerge in perception (segregation). Sussman and Winkler (2001) postulated that the dynamic process of sensory input updating continuously in the auditory system.

Note also that source reconstruction studies (Opitz et al., 2002; Doeller et al., 2003) suggest that MMN is generated bilaterally in the temporal cortex. These previous studies used dipole modeling over different time windows to explain the topographic distribution of MMN. They observed that the early component (in the 90-120 ms time window) originates in the superior temporal gyrus (STG) while the late component (in the 140-170 ms time window) was localized to the left and right inferior frontal gyrus (IFG). This result suggests that the temporal area might involve processing the signal physical properties, whereas the frontal area might reflect interpretation sensory input of signal processing.

In addition, from the model adjustment perspective, MMN can be used to study how the brain detects deviant input and modulates its processing by top-down mechanisms. A previous study by Sussman et al., (2003) supported this hypothesis by providing evidence for the top-down modulation of the deviance detection system. The results of behavioral and electrophysiological analyses suggested that the predictability of irrelevant sound changes was related to involuntary switches of attention. When irrelevant sound changes occurred unpredictably, reaction times were longer and attention-related ERP components (P3 peak) could be observed. However, when sound changes occurred in a predictable fashion, the correlates of involuntary attention switches were not observed. The results are consistent with the notion of top-down control over stimulus-driven switching of attention.

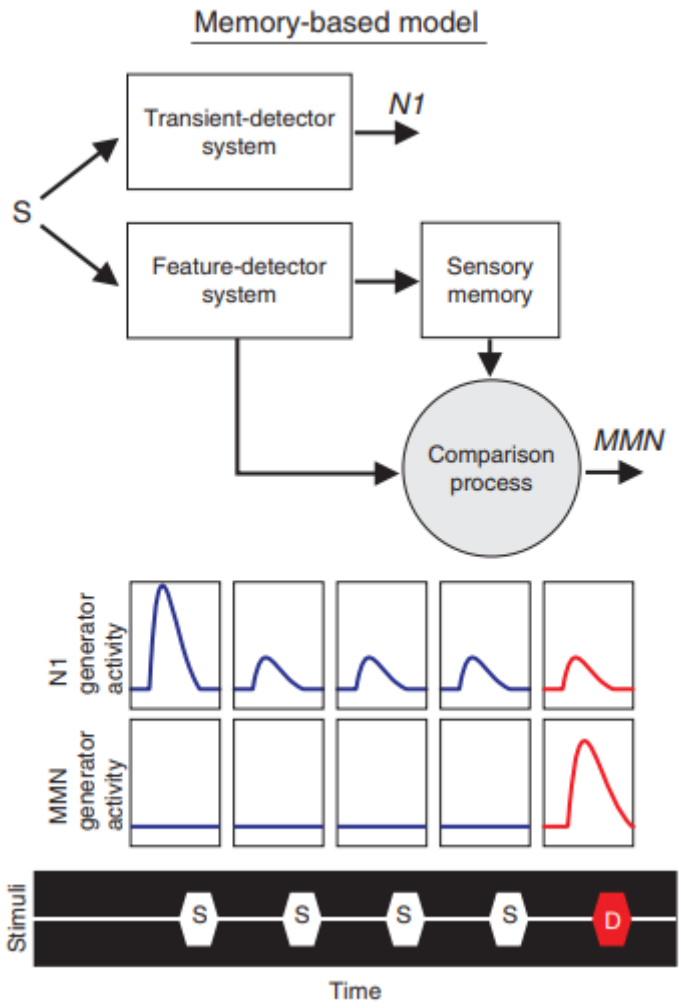

Figure 5.1. Model Adjustment

According to the model adjustment hypothesis, the N1 component of the response to an unexpected stimulus reflects an analysis by a transient detector system. The results of the analysis are passed onto a sensory memory system. Based on the comparison between past stimuli in sensory memory and new stimuli, the MMN response reflects the difference between them. The bottom panels are depicting N1 generator activity and MMN generator activity. Four standards (S) are followed by one deviant (D) and the corresponding neural responses are shown in blue and red lines. The first N1 response is larger than the MMN response. In contrast, the MMN generator response is only visible following the deviant stimulus (May and Tiitinen, 2010).

1.4.2 Adaptation hypothesis

Unlike the model adjustment hypothesis, the adaptation hypothesis supposes that the attenuated response following standard stimuli could be simply due to neural adaptation (Figure 5.2). A previous study by May and Tiitinen (2010) suggested that the attenuation response may be due to synaptic depression and lateral inhibition mechanisms. According to the adaptation hypothesis, the N1 response, which delays or suppresses the activation of calcium-dependent potassium channels, leads to a slow hyperpolarizing current and decreases neuronal firing rates. Thus, the neural adaptation results from pre-synaptic mechanisms and rests on post-synaptic responsiveness changes (Faber and Sah, 2003). The adaptation hypothesis encompasses the SSA, which in this view is due to fast time constants of adaptation during the stimulus sequence, and demonstrates a simple, low-level mechanism mimicking more sophisticated perceptual-cognitive effects (Ulanovsky et al., 2004b; Costa-Faidella et al., 2011).

However, while both adaptation and adjustment hypotheses have been used to interpret the MMN, both hypotheses have shortcomings in explaining this phenomenon. The adjustment hypothesis is weakened by the lack of clarity regarding the extent to which MMN reflects a pure measure of deviance; the absence of direct evidence for a population of neurons capable of higher-order change detection; and the lack of consistent support from animal and intracranial studies (Edwards et al., 2005; Taaseh et al., 2011; Fishman and Steinschneider, 2012; Harms et al., 2014). The adaptation hypothesis has failed to account for the large MMN elicited by repetition deviants (Horvath and Winkler, 2004; Macdonald and Campbell, 2011; Wacongne et al., 2012), stimulus omissions (Hughes et al., 2001; Wacongne et al., 2011; Salisbury, 2012), and unpredicted vs. predicted deviant tones (Sussman et al., 1998; Sussman and Gumenyuk, 2005).

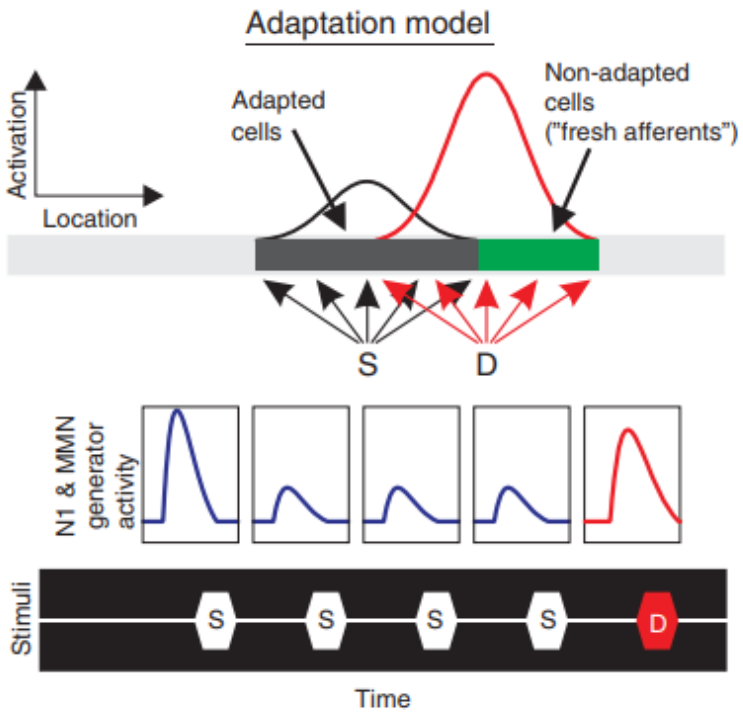

Figure 5.2. Adaptation Model
In the adaptation model, when stimuli are presented, the activations of adapting cells and non-adapting cells are overlapping. The standard (repetitive) stimuli lead to cell tuning, resulting in adaption and suppression. In contrast, when deviant (novel) stimuli are presented, the responses of non-adapting cells increase. Unlike the model adjustment hypothesis, this model suggested that the N1 and MMN are generated by the same neural populations (May and Tiitinen, 2010).

1.4.3 Predictive coding

A new perspective, which integrates model adjustment and adaptation, has been proposed to account for the neurobiological mechanisms of deviance. This framework is referred to as predictive coding. This is one of the most comprehensive neural function theories that account for how the brain perceives the world (Heilbron and Chait, 2018). Consistent with the model-adjustment hypothesis, in this view, the input stimuli are compared with information inferred from past events through an internal model, and the following error is reduced by continuously updating

the model. However, when predictive coding is used to explain findings based on oddball paradigms, it faces a major methodological limitation as it confounds adaptation and expectation effects (Ruhnau et al., 2012). In an oddball sequence, the adaptation effect induced by repetitive tones is known to cause a reduced auditory N1. By subtracting the standard response (including N1 peaking at 100 ms) from the deviant response (peaking at around 100-200 ms), MMN could be contaminated by the N1 and overestimated. However, a suitable control sequence can help estimate adaptation and deviance detection separately. For example, in the many-standard sequence, the target tone is embedded within a sequence of other random tones. This sequence does not generate repetition suppression in response to the target tone, while controlling for the state of refractoriness of the auditory system (Schröger and Wolff, 1996). This comparison between the response to the target tone in the control sequence and the standard-evoked response accounts for repetition suppression while the difference between the deviant-evoked response and the target response can be imputed to prediction error (Figure 5.3).

Thus, predictive coding establishes a computational model of neuronal activity via a Bayesian inference system (Friston, 2005), which involves estimating the posterior probability based on a prior probability distribution. In this view, the MMN reflects the discrepancy between sensory information from the environment and predictions based on an internal model of the environment (Auksztulewicz and Friston, 2016).

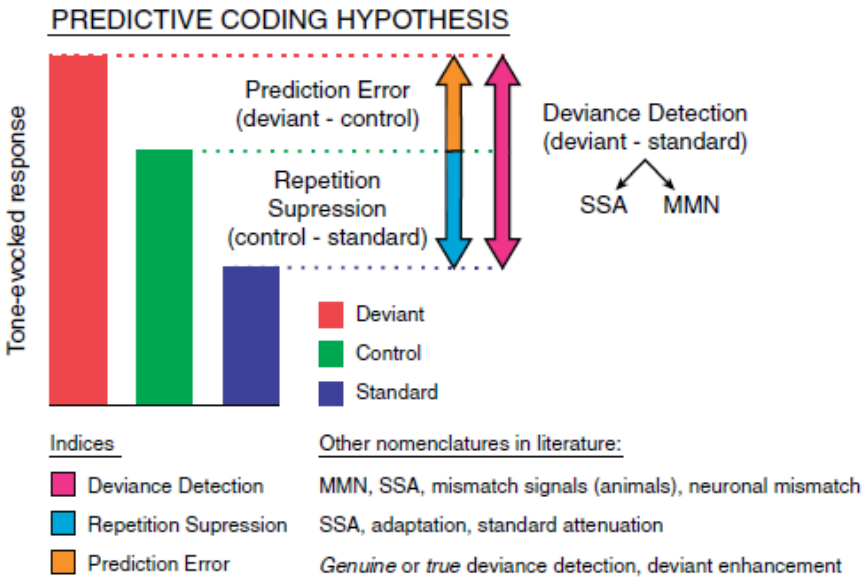

Figure 5.3. Predictive Coding
Based on the predictive coding theory, repetition suppression occurs when the control-standard comparison has a positive value, and the prediction error is observed when the deviant–control comparison has a positive value. Deviance detection indicates a more general concept that is equivalent to the deviant-standard comparison (Carbajal and Malmierca, 2018).

1.5 Deviance detection across acoustic features

Previous research has attempted to explain the mechanisms of deviance detection between the incoming sound and past sounds based on various auditory features (e.g., sound duration, fundamental frequency, vowel formant, spatial location, etc.) This chapter will review MMN studies in animal and human models using different deviant features and paradigms (Table 1).

1.5.1 Sensitivity to deviance in duration

The time dimension plays a vital role in the generation of MMN. A previous study (Umbricht et al., 2005) investigated event-related potentials (ERP) evoked by deviant stimuli in different auditory stimulation paradigms in mice to determine if duration MMN can be observed in an animal model. It has been found that standard-deviant difference occurred on a 50% reduction

(from 100 to 50 ms) in duration. Such results would be predicted based on the different onsets of the stimulus differences and indicated that MMN shifts according to the onset of the detectable difference between deviant and standard stimuli. In another study (Roger, Hasbroucq, Rabat, Vidal, & Burle, 2009), the authors investigated the discrimination between the standard and the deviant by recording electroencephalographic responses in the rat. The results showed that MMN was observed in response to 33% (100 ms deviant vs. 150 ms standard) but not 16% changes in duration. Recently, Lipponen et al. (2019) investigated the detection of vowel duration changes in anesthetized mice. Epidural recordings of auditory-evoked potentials were obtained in anesthetized mice exposed to change in sound duration from 100 to 200 ms in 10 ms steps. They observed a differential response only to the shortest (110 ms) deviant sound interspersed with the 200 ms standard sound (45% duration decrease).

In contrast to rodent studies, in humans (Joutsiniemi et al., 1998) MMN could also be elicited by a shorter duration tone (25 and 50 ms) than the 75 ms standard tone. Although the MMN peak amplitude for 25 ms deviants was larger than for 50 ms deviants, no significant difference was observed in peak latencies. Robust MMN has been detected following duration increments and decrements even for 10% changes in sound duration (100 ms vs. 110 ms) (Jaramillo et al., 2000). Thus, the neurophysiological detection of sound duration changes in rodents seems not to be as advanced as in humans. This suggests that deviant detection might depend on the complexity of the network of the brain.

1.5.2 Sensitivity to deviance in frequency

Previous reports utilizing epidural recordings in the rat have found evidence of MMN-like activity to pure-tone frequency deviants. Ruusuvirta et al.(1998) recorded ERP from the auditory cortex in

anesthetized rats to different tones (2000 / 2500 Hz). The MMN-like activity was observed as a significant difference between standards and deviants, 63-253 ms from stimulus onset. This result implied that deviant tones were neurophysiologically discriminated from standard tones in anesthetized rats. Another study (Astikainen et al., 2011) recorded activity over the primary auditory cortex in urethane anesthetized rats. In oddball conditions, tone frequency was manipulated to produce a small difference of 5% (200 Hz deviants) and a large difference of 12.5% (500 Hz deviants). Mismatch responses were observed at 60-100 ms after stimulus onset for a frequency increase of 5% and 12.5%, but not for similarly descending deviants.

Consistent with the animal models, a previous study in humans (Sams et al., 1985) observed the MMN in auditory frequency discrimination. They found that the MMN is much more pronounced following larger differences in tone frequency. Consistent with the previous study, two recent studies have analyzed not only frequency but also spectral information that facilitates pitch processing (Tervaniemi et al., 2000; Grimm et al., 2011). Thus, both animal and human studies strongly support the notion of a hierarchically organized novelty and deviance detection in the auditory system.

1.5.3 Sensitivity to deviance in phonetic features of vocal sounds

The ability to represent speech sounds is a necessary condition for understanding speech. Interestingly, this ability is not unique to humans but can also be observed in animal models. In a previous study in guinea pigs (Kraus et al., 1994), mismatch responses were observed in the auditory thalamus and cortex. The authors found that the /ba/-/wa/ pair elicited a strong mismatch response, whereas /ga/-/da/ mismatch responses did not occur in the thalamus. Moreover, mismatch responses were observed only in the non-primary areas of the auditory pathway. These

results indicate that different speech features may be processed differently and that certain acoustic features require processing at the cortical level. A recent study (Ahmed et al., 2011) presented synthesized spoken syllables (/da/, /ga/, and /ba/) to anesthetized rats and recorded their neural activity in the primary auditory cortex. The authors found that evoked responses had a higher amplitude to the deviant /ba/ than to the standard /ba/ in the oddball condition. The findings suggested that animal brains can represent human speech change detection mechanisms analogous to those underlying the MMN in humans.

In humans, neural representations of speech sounds are already established pre-attentively. Previous studies (Näätänen et al., 1997; Sharma and Dorman, 2000) found that the brain's automatic change-detection response differed depending on the listener's native language. A recent study (Altmann et al., 2014b) tested the neurophysiological effects of speech change detection using magnetoencephalography (MEG). The authors used manipulated consonant-vowel (CV) sounds, including /ba/, /da/, /bo/, and /do/, and morphed either consonants (e.g., /ba/ - /da/) or vowels (e.g., /ba/ - /bo/) from 0% to 100% in 10% steps. Behaviorally, they found that subjects showed more pronounced effects for the morphed consonants compared to vowels. They also observed a difference in representations between consonants and vowels at the neural level, in which consonants revealed more pronounced effects in the left superior temporal gyrus (STG) and sulcus. Thus, MMN can also reflect a processing level in language-specific categories and anatomical locations for representations of phonemic deviance.

1.5.4 Sensitivity to deviance in spatial location

An ability to judge the spatial localization of sound sources is a skill of considerable adaptive value (Darwin and Hukin, 2000). Two primary sound localization cues, i.e., interaural time difference

(ITD) and interaural level difference (ILD), exhibit anatomical separation already in the brain stem (Grothe et al., 2010). Recent studies have attempted to measure the neuronal response in different cortical fields across sound source locations (Recanzone, 2000; Stecker et al., 2003; Harrington et al., 2008). In one of the previous studies (Recanzone, 2000), the authors compared the cortico-cortical and cortico-thalamic connections with respect to their spatial tuning. The results have shown that neurons in the caudomedial fields better predict sound localization than the primary auditory cortex. In the auditory cortex, other studies (Stecker et al., 2003; Harrington et al., 2008) also observed that the posterior auditory field (PAF) units responded with longer latency and stronger stimulus location compared to the anterior auditory field (AAF) and primary (A1). This evidence supports the notion that the response properties of neurons in different cortical areas are functionally distinct with respect to the spatial localization of sounds.

Another study (Roger et al., 2009b) measured the mismatch response to temporal cue discrimination to study central auditory mechanisms underlying spatial processing. The authors found that a difference between deviants and standards was observed for all temporal difference conditions except the smallest difference (16% change). Overall, the topographical and electrophysiological results support the hypothesis that ILD and ITD cues are processed by distinct cortical processing.

However, it should be noted that other invasive studies have found evidence to the contrary – namely that specialization of cortical regions for spatial processing is at best partial. For instance, in the auditory cortex of ferrets, spatial location has been found to be processed in a distributed manner, together with other acoustic features such as pitch and timbre (Bizley et al., 2009).

Similarly, distributed processing of auditory spatial information has been found in the cat auditory cortex (Stecker et al., 2003; Furukawa et al., 2000).

In humans, a previous electroencephalography (EEG) study (Schröger, 1996) has shown a significant difference in the latency, amplitude, and topographies of MMN to ILD and ITD comparable to the sum of MMN response to single deviants (ITD or ILD). Other studies (Altmann et al., 2014a; Altmann et al., 2017) also support the notion of independent processing of ITD and ILD cues at the level of the MMN along with the difference between the frequency ranges (low and high). These findings suggest that the additivity of ITD and ILD is processed independently in both animal and human studies.

Table 1. Deviance detection studies using manipulations of different acoustic features

Publication	Species	Recoding	Probability	Deviant	Stimulus Characteristics		Comments
			(%)		Standard	Deviant	
Umbricht et al., (2005)	Mice	Scalp EEG	10-90	Duration & Frequency	10kHz 100ms	5-14kHz 50, 250ms	Duration deviance paradigm showed typical human MMN response.
Roger et al., (2009)	Long Evans rats	Scalp EEG	16.67-66.67	Duration	3kHz 150ms	50,75,100 and 125ms	MMN response showed a discrimination threshold between 16% and 33% change.
Lipponen et al., (2019)	Mice	Epidural	10-90	Duration	/a/ speech 200ms	110-180ms Steps 10ms	The shortest deviant (110ms) showed a significant difference.
Joutsiniemi et al., (1998)	Human	EEG	6-94	Duration	700Hz 75ms	700Hz 25 and 50ms	Both shorter tones showed MMN response.
Jaramillo et al., (2000)	Human	EEG	20-80	Duration	White noise 100ms	Increment 110, 150 or 200ms Decrement 1, 10 or 50ms	MMN was elicited by both stimulus decrement and increments.
Ruusuvirta et al.(1998)	Wistar rats	Epidural	5-95	Frequency	2000Hz 50ms	2500Hz 50ms	ERPs deviant tone showed a significantly different from the standard tone at 63-243 ms
Astikainen et al., (2011)	Spraque Dawley rats	Epidural	10-90	Frequency	4000Hz 50ms	Small:3800-4200Hz Large: 3500-4500Hz	Increment frequency deviant (small and large) showed mismatch response at 60-100ms but not for similar decrement.
Sams et al., (1985)	Human	EEG	20-80	Frequency	1000Hz 50ms	1002, 1004, 1008, 1016 or 1032Hz	MMN was elicited by 1016 and 1032 Hz, and 1008 Hz tended small MMN.
Grimm et al., (2011)	Human	EEG	20-80	Frequency	Low:1200Hz High:2580Hz	Low: 800 Hz Hight: 3780Hz	The low-frequency condition showed more negative mean amplitude than the high condition.
Tervaniemi et al., (2000)	Human	EEG	10-90	Frequency	3paris:500-1500Hz 5paris: 500-2500Hz	3pairs: 513,1026 and 1539 5paris: 513,1026, 1539, 2052 and 2565 Hz	MMN amplitude enhanced rich sounds compared with sinusoidal tone.

Kraus et al., (1994)	Guinea pig	Epidural	10-90	Consonant	/ga/ /ba/	/da/ /wa/	At thalamic level, /ba/-/wa/ showed mismatch response. Moreover, mismatch response observed non-primary auditory pathway.
Ahmed et al., (2011)	Spraque Dawley rats	Epidural	10-90	Consonant	/da/	/ga/ or /ba/	Deviant /ba/ showed higher amplitude than standard /da/.
Näätänen et al., (1997)	Human	EEG	15-85	Vowel	Finnish and Esotonian /e/	/6/ and /o/	The vowel prototype of the native language presented a larger MMN.
Sharma and Dorman, (2000)	Human	AEP	15-85	Consonant	/ba/-/pa/ change voice on set (VOT) 10ms	/ba/-/pa/ change voice on set (VOT) 50ms	Robust MMN was seen only in Hindi listeners (non-native) and not in English listeners (native).
Altmann et al., (2014)	Human	MEG	Behanvioral test	Consonant and vowel	/da/ -/ba/, /ba/-/bo/, /bo/-/do/, and /da/-/do/ for each morph 10% steps change		In the behavioral test, consonants have a more substantial categorical effect than vowels. At a neural level, consonants revealed a categorical effect in the left STG.
Roger et al., (2009)	Long Evans rats	Epidural	30-70	ITD	150ms	125, 100, 75 and 50ms	The difference between deviants and standard was significant for 50, 75, and 100 ms but not 125 ms deviant.
Schröger (1996)	Human	AEP	12-88	ITD/ILD	ITD: 0 µs ILD: 0 dB SPL	ITD: delay 300 µs ILD: 11dB SPL	The MMN obtained with ILD/ITD deviants were larger than one location cue only.
Altmann et al., (2014, 2017)	Human	EEG	1.67-86.67	ITD/ILD	ITD: 0 µs ILD: 45 dB SL	ITD: 200 or -200 µs ILD:average for two ears on dB scale	Significant MMN for ILD/ITD deviants in low and high frequency, but an incongruent combination ILD/ITD was only detectable in a lower frequency range.

1.6 Memory for auditory information

Auditory sensory memory is a prerequisite for MMN, as in the process of deviance detection, the auditory system compares the input from the deviant auditory information with the sensory memory representations based on the preceding auditory stimulation (Näätänen and Michie, 1979). In a recent study (Haenschel et al., 2005), it has been found that ERP changes correlated with the repetition of the standard stimulus observed as a positive polarity between 50 and 250 ms post-stimulus, which is termed repetition positivity (RP). This RP was recorded from the frontal electrode when participants listened in passive or active conditions. These RP effects are similar to adaptation effects in the primary auditory cortex neurons to the repetition of sounds, namely the SSA (Ulanovsky et al., 2003). Both studies reported their findings to occur without overt attention to sounds and to develop rapidly. An earlier study (Näätänen and Rinne, 2002) demonstrated stimulus repetition during auditory memory-trace formation. The authors exposed participants to randomized sequences of stimuli (ten different tones or complex frequency-glide stimuli). They found that repetition negativity response was elicited only by stimulus repetitions. These results indicated that the repeated experience with the same sensory input stimulus leads to suppressing the neuronal response. This particular repetition effect was mediated by a change in synaptic plasticity for a memory trace.

Despite the efforts of previous studies, the neural mechanisms underlying echoic memory traces are still not fully understood. One of the caveats of these studies is that the constant presentation of sounds reflects realistic auditory scenes only to a limited extent. In realistic environments, listeners must learn from complex sounds and associate them with various sound sources. Recent studies (Andrillon et al., 2015; Barascud et al., 2016; Southwell and Chait, 2018) investigated how human listeners employ auditory memory processing to discover statistical structure in complex

sound sequences. Andrillon et al. (2015) applied psychophysics, EEG, and modeling to the noise learning paradigm, in which participants are exposed to reoccurring identical noise snippets, interspersed with other noise snippets. They found that repeated exposure to the same noise snippet is accompanied by a rapid formation of behavioral and neural selectivity to these stimuli. More specifically, in a within-trial repetition detection task, participants showed faster reaction times and higher accuracy for noise stimuli that also reoccurred across trials than for those stimuli that were presented anew in each trial. In the EEG analysis, the inter-trial phase coherence (ITPC) increased for reoccurring noise stimuli compared to other conditions. Furthermore, the EEG signals showed an ERP-like pattern locked to noise reoccurrence within trials. The amplitude of this ERP correlated with the behavioral performance in the within-trial repetition detection task, indicating that this neural response induced by learning new sounds has behavioral significance and can be interpreted as memory-evoked potentials (MEPs). These MEPs could be source-localized to higher-order auditory regions, which had previously been linked to auditory mismatch responses (Berti et al., 2000; Näätänen et al., 2005).

Another study (Barascud et al., 2016) has investigated how listeners are sensitive to the emergence of complex patterns within rapidly evolving sound sequences. While the main finding of the study was that acoustic pattern regularity correlated with the amplitude of sustained neural activity, the authors also found that an MMN-like response was evoked by the transition from regular to random pattern conditions. This result is consistent with the notion that the brain continually accumulated the internal model, and this model is shaped by both the memory of statistical regularities in the stimulus sequence and the incoming information processing.

Similarly, Southwell and Chait (2018) investigated whether the regularity of stimulus sequences influences the neural responses to deviant tones presented in these sequences. The study found that the neural responses to deviant (outlier) tones were greater when the outliers were embedded in regular sequences than in random sequences. Interestingly, a source reconstruction analysis indicated that the increased deviance response in regular sequences could be localized to the temporal and orbitofrontal regions, which have been proposed to exert top-down modulation of the auditory cortex. These findings provide converging evidence for an interplay between deviance detection and continuous monitoring of the sequence structure, which enables memory formation.

1.7 Thesis overview

This thesis investigates the neural correlations of prediction violation based on multiple acoustic features and implicit learning based on complex acoustic stimuli. In particular, to assess whether the neural correlates of deviance detection across various stimulus features show evolutionary conservation, I recorded electrophysiological signals in different species. Experimental Chapters 2-4 summarise the investigations into the research questions that I have undertaken throughout my Ph.D studies.

Chapter 2: Cortical mapping of mismatch responses to independent acoustic features

This chapter analyses the mismatch response amplitudes, topographies, and latencies following prediction violations along several acoustic features (duration, pitch, interaural level difference, or consonant identity). To quantify mismatch responses, I recorded auditory cortical activity with a 64-channel ECoG array in 9 adult female Wistar rats. I found that mismatch responses differ

significantly in terms of their spatial distribution from the sound-evoked response. However, there were also considerable differences between individual rats regarding the spatial distribution of mismatch responses to prediction violations along different acoustic features. This study suggested that different but largely idiosyncratic neural populations mediate the predictive processing of different stimulus features.

Chapter 3: Do auditory mismatch responses differ between acoustic features?

As an extension of the previous chapter, testing the evolutionary conservation of domain specificity of mismatch responses, I measured MMN in humans and tested whether it differentiates between prediction violations along different features. To this end, EEG signals were recorded in normal-hearing participants. Using a multivariate decoding analysis, I found that acoustic features could be decoded at later latencies than typical for MMN. This finding indicates that the process of deviance feature detection might be separate from general mismatch detection.

Chapter 4: Neural correlates of auditory pattern learning in the auditory cortex

This chapter characterizes the neural correlates of auditory pattern learning in the auditory cortex based on repetitive exposure to a specific sound sequence. In this study, I recorded neural activity in the auditory cortex using ECoG in 9 anesthetized young adult female Wistar rats. I found that auditory cortical activity in the beta frequency band significantly decreased for repetitive sequences relative to fresh sequences. This suggests that neural correlates of auditory learning formation can be observed in animal models even under anesthesia.

The thesis entitled 'Predictive processing of multiple acoustic features in the auditory cortex' presented by HyunJung An to obtain the degree of PhD in Neuroscience corresponds to a compendium of scientific articles already published or accepted for publication. The articles are listed below.

Study I

Cortical mapping of mismatch responses to independent acoustic features

Authors: HyunJung An[1], Ryszard Auksztulewicz[1,2], HiJee Kang[1], Jan W.H. Schnupp[1]

Affiliation: [1] Department of Biomedical Sciences, City University of Hong Kong, Hong Kong

[2] Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

Journal: Hearing Research

DOI: https://doi.org/10.1016/j.heares.2020.107894

Study II

Do auditory mismatch responses differ between acoustic features?

Authors: HyunJung An[1], Shing Ho Kei[1], Ryszard Auksztulewicz[1,2]*, Jan W.H. Schnupp[1]*

* Equal contribution

Affiliation: [1] Department of Biomedical Sciences, City University of Hong Kong, Hong Kong

[2] Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

Journal: Frontiers in Human Neuroscience

DOI: https://doi.org/10.3389/fnhum.2021.613903

Study III

Neural correlates of auditory pattern learning in the auditory cortex

Authors: HiJee Kang[1], Ryszard Auksztulewicz[1,2], HyunJung An[1], Nicolas Abichacra[1], Mitchell Sutter[3], Jan W.H. Schnupp[1]*

Affiliation: [1] Department of Biomedical Sciences, City University of Hong Kong, Hong Kong

[2] Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

[3] Center for Neuroscience, and Section of Neurobiology, Physiology and Behavior, University of California, Davis, California

Journal: Frontiers in Neuroscience

DOI: https://doi: 10.3389/fnins.2021.610978

Chapter 2.

Cortical mapping of mismatch responses to independent acoustic features

HyunJung An[1], Ryszard Auksztulewicz[1,2], HiJee Kang[1], Jan W. Schnupp[1]

[1] Department of Biomedical Science City University of Hong Kong, Hong Kong

[2] Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

Address for correspondence: Prof. Jan W.Schnupp.
City University of Hong Kong
Dept. of Biomedical Sciences, 31 To Yuen Street, Kowloon Tong
Hong Kong
Tel +852 xxxxxxxx
Fax +852 34420549
E-mail xxxxxxxxx@cityu.edu.hk

Word counts

Number of words in abstract: 343
Number of words in the manuscript (excluding references and abstract): 5174
Number of tables/figures: 1 tables, 5 figures
Number of references: 59

Keywords: Predictive coding, Mismatch negativity (MMN), stimulus-specific adaptation (SSA), Multiple sensory features, Electrocorticography (ECoG), Auditory event-related potential (ERP)

Abstract

Predictive coding is an influential theory of neural processing underlying perceptual inference. However, it is unknown to what extent prediction violations of different sensory features are mediated in different regions in auditory cortex, with different dynamics, and by different mechanisms. This study investigates the neural responses to synthesized acoustic syllables, which could be expected or unexpected, along several features. By using electrocorticography (ECoG) in rat auditory cortex (subjects: adult female Wistar rats with normal hearing), we aimed at mapping regional differences in mismatch responses to different stimulus features.

Continuous streams of morphed syllables formed roving oddball sequences in which each stimulus was repeated several times (thereby forming a standard) and subsequently replaced with a deviant stimulus which differed from the standard along one of several acoustic features: duration, pitch, interaural level differences (ILD), or consonant identity. Each of these features could assume one of several different levels, and the resulting change from standard to deviant could be larger or smaller. The deviant stimuli were then repeated to form new standards. We analyzed responses to the first repetition of a new stimulus (deviant) and its last repetition in a stimulus train (standard). For the ECoG recording, we implanted urethane-anaesthetized rats with 8×8 surface electrode arrays covering a 3×3 mm cortical patch encompassing primary and higher-order auditory cortex.

We identified the response topographies and latencies of population activity evoked by acoustic stimuli in the rat auditory regions, and mapped their sensitivity to expectation violations along different acoustic features. For all features, the responses to deviant stimuli increased in amplitude relative to responses to standard stimuli. Deviance magnitude did not further modulate these mismatch responses. Mismatch responses to different feature violations showed a heterogeneous

distribution across cortical areas, with no evidence for systematic topographic gradients for any of the tested features. However, within rats, the spatial distribution of mismatch responses varied more between features than the spatial distribution of tone-evoked responses. This result supports the notion that prediction error signaling along different stimulus features is subserved by different cortical populations, albeit with substantial heterogeneity across individuals.

1. Introduction

Predictive coding is an influential theory of neural processing underlying perceptual inference. Successful listening requires not only tracking stable auditory objects, but also detecting unexpected changes to the auditory scene. Mismatch negativity (MMN) is one of the most well-studied physiological correlates of deviance detection in the auditory system. When presented with unexpected input, brain activity recorded using electroencephalography in humans exhibits a well-described pattern of neural response waveforms showing an enhanced response to novel stimuli. Conversely, repetitive, and therefore expected, stimulation leads to reduced responses. The difference in response waveforms between responses to unexpected as opposed to expected stimuli is known as the MMN, and it is well defined based on previous studies in humans (Näätänen, 1990).

MMN is a macroscopic signature of deviance detection derived from bulk measures of neural activity such as electroencephalogram (EEG). Meso- and microscopic correlates of mismatch responses have also been described at the level of small neuronal populations (multi-units) and single neurons along the auditory pathway, showing an increased response amplitude to a novel stimulus (oddball) relative not only to a repeated standard stimulus (Malmierca et al., 2009), but also to a control stimulus that is predictable but not repeated (Nieto-Diego and Malmierca, 2016), suggesting true deviance detection. This particular type of adaptation, known as stimulus-specific adaptation (SSA), quantifies the changing neuronal firing rates to a deviant stimulus compared with a standard. SSA can be observed in both primary and secondary areas of the auditory cortex (Ulanovsky et al., 2003; Malmierca et al., 2009; von der Behrens et al., 2009; Antunes et al., 2010; Richardson et al., 2013; Duque et al., 2016; Nieto-Diego and Malmierca, 2016; Parras et al., 2017).

SSA responses are thought to reflect deviance detection mechanisms at the microscopic level, and they have been suggested to form a counterpart of the MMN which is a more macroscopic response detectable on the scalp (Ulanovsky et al., 2003; Escera and Malmierca, 2014; Nieto-Diego and Malmierca, 2016).

Traditionally, two main, not necessarily mutually exclusive, hypotheses have been put forward to interpret these physiological observations. According to the "model adjustment hypothesis" (Saarinen et al., 1992), the MMN reflects an online update of the prior perceptual model, previously established due to sensory experience, by comparing it with actual auditory input. Meanwhile, the adaptation hypothesis (May et al., 1999) claims that the attenuation of the response to standard stimuli could be simply due to neural adaptation. More recently, these hypotheses have been united under the predictive coding framework, according to which the brain entails and continuously revises an internal model of the world. In this theory, the model updates are based on prediction error signals, which result from comparing the internal model predictions with the actual sensory inputs (Friston, 2005; Bastos et al., 2012; Auksztulewicz and Friston, 2016). When sensory information reaches a given stage of sensory or cortical processing, it is suppressed by predictions signalled from higher-order stages in a descending (top-down) manner. If new sensory information cannot be predicted by the brain's internal model, the resulting prediction errors are propagated in a bottom-up (ascending) manner to higher-order regions and update the subsequent predictions. Hence, lower and higher cognitive stages keep communicating through reciprocal pathways, with descending connections mediating prediction signalling, and ascending connections mediating prediction error signalling. This results in a gradual suppression of prediction error signalling (model adjustment) and a rapid decrease of synaptic efficacy (adaptation) when a deviant stimulus is first encountered and then repeated to form a standard. Specifically, according to the model

adjustment hypothesis, the MMN is thought to reflect an online update of the prior perceptual model, based on a comparison between previous sensory experience and the actual auditory input (Saarinen et al., 1992). Unlike the model adjustment hypothesis, the adaptation hypothesis claims that the attenuation of the response to standard stimuli could be simply due to neural adaptation (May et al., 1999; Jaaskelainen et al., 2004; May and Tiitinen, 2010).

While prediction error suppression is a widely accepted theory explaining mismatch responses, it has been suggested that predictions and prediction errors corresponding to different stimulus features are mediated by different neural populations, and even different mechanisms (Auksztulewicz et al., 2018a; Stefanics et al., 2019). On the other hand, in animal studies, although mismatch responses have been shown to occur following changes in repetitive syllable patterns (Ahmed et al., 2011; Mahmoudzadeh et al., 2017), syllable duration (Lipponen et al., 2019), and acoustic frequency (Eriksson and Villa, 2005), no robust regional differences to various stimulus features have been identified. Directly comparing results of studies in animal models and human volunteers is challenging because of methodological differences: while most animal studies apply a classical MMN design with a low constant probability (10%) of occurrence of a deviant and a high probability (90%) of the standard tone (Javitt et al., 1996; Pincze et al., 2001; Ulanovsky et al., 2003; Ulanovsky et al., 2004b; Roger et al., 2009b), human studies often use more complex paradigms with dynamically evolving (and occasionally switching) probabilities of deviant stimuli (Baldeweg et al., 2004; Haenschel et al., 2005; Garrido et al., 2009b).

Here, building upon previous studies that characterized the spatial distribution of mismatch responses to a single auditory feature (Nieto-Diego and Malmierca, 2016), we aim to broaden our

understanding of predictive coding in auditory regional mapping differences across multiple acoustic features.

2. Material and methods

All experimental procedures were approved by the Committee on the Use and Care of Animals at City University of Hong Kong and under license by the Department of Health of Hong Kong [Ref. No. (17-72) in DH/SHS/8/2/5 Pt.1].

2.1 Subjects and surgical procedures

The subjects were nine female adult Wistar rats acquired from the Chinese University of Hong Kong. The rats were between 8 and 20 weeks of age (mean = 10 weeks) and weighed 200 - 310 g (mean = 238 g) at the time of the experiment. All rats were normal hearing (click ABR thresholds < 20 dB) with no prior exposure to the stimulus sequences described below. At the start of each recording experiment, we used mixture of ketamine (80 mg/kg, Intraperitoneal injection; i.p) and xylazine (12 mg/kg, i.p) to induce anaesthesia and received an injection of the anti-inflammatory dexamethasone (0.2 mg/kg, i.p) before surgery. Body temperature was maintained with a heating pad at 36° ± 1 C. This was replaced by urethane administration during subsequent recording. Ketamine + xylazine was chosen as the main surgery anaesthetic as it takes less time to induce deep anaesthesia than urethane. Urethane (0.75mg/kg, i.p) was administered about one hour after induction with the ketamine + xylazine. Therefore, ketamine + xylazine should not affect the neural activity recorded during urethane anaesthesia. This protocol is based on previous studies in rodents (Malmierca et al., 2019). The depth of anesthesia was controlled by regular testing of the toe pinch and the absence of withdrawal reflex. If required, extra doses (0.1- 0.2 ml) of urethane were administered. The anesthetized animal was placed in a stereotaxic frame in which hollow ear

bars were set to deliver sound and fix head for craniotomy. The skin and muscle tissue over the right temporal side of the skull were removed. A unilateral craniotomy was performed to expose a 5 mm × 4 mm region over the right primary auditory cortex, as shown in Figure 1 (2.5 mm posterior from the bregma, and ventral from the temporal edge of the lateral skull surface) (Doron et al., 2002; Lamas et al., 2017).

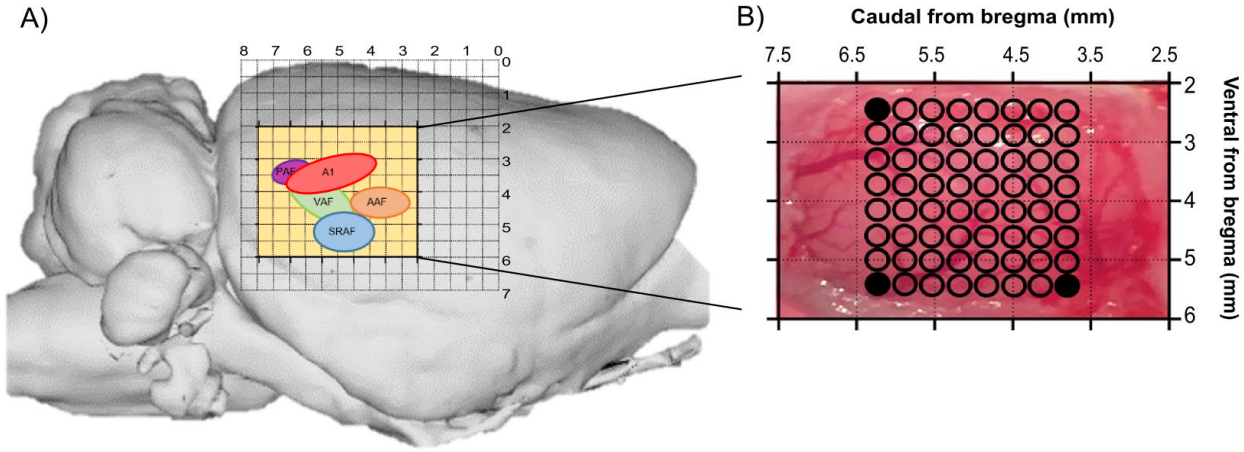

Figure 1. Experiment setup.

A) Rat brain atlas template: 8-7 mm grid projected onto a lateral view of the right hemisphere of the rat brain. Horizontal and vertical numeric values represented distance with respect to bregma (0,0) (Bakker et al., 2015). Yellow: Schematic craniotomy site including the 5 auditory cortical fields; A1(primary auditory cortex), VAF(ventral auditory field), SRAF(suprahinal auditory field), AAF(anterior auditory field), PAF(posterior auditory field) (Polley et al., 2007). *B) Photograph of a craniotomy, with superimposed drawing of the ECoG recording sites for one animal. Empty circles show the locations of the 61 ECoG electrode sites, while filled circles represent 3 reference electrodes.*

2.2 Auditory paradigm

In this experiment, two consonant-vowel (CV) stimuli /da/ and /ba/ were selected from a corpus of natural speech syllables (Ives et al., 2005) and digitally edited using STRAIGHT software toolbox (Kawahara, 2006) for Matlab R2018b (Mathworks Inc., Natick, USA) to match the raw syllables for vowel duration and fundamental frequency (F0). We then used STRAIGHT to obtain stimulus tokens varying systematically in vowel duration, fundamental frequency (F0), inter-aural level differences (ILDs), and the onset consonant (/da/-/ba/, obtained by mixing the spectrograms corresponding to the two syllables at different ratios). In this study, stimulus features are largely based on previous animal studies which have investigated mismatch responses to changes in duration, pitch, ILD, or consonant (Roger et al., 2009b; Walker et al., 2009; Ahmed et al., 2011; Li et al., 2019). Since the previous animal studies did not include auditory cortical mapping across different features, we selected these features for our study. There is also a precedent in the human literature (Phillips et al., 2015), which compared amplitudes and sources of mismatch responses to duration, pitch/frequency, ILD/location, as well as intensity and gap. Our initial rationale to replace e.g. intensity with another condition (consonant) was that we wanted a relatively small set of features that would have the potential to map onto distinct auditory regions; since intensity is known to modulate processing at all stages from the cochlea, we reasoned that intensity mismatch is likely to have less localised effects than other features. This set of syllables was then concatenated to generate stimulus sequences containing syllables that could vary along one of the four different features: duration, pitch, ILD, or consonant. The features were manipulated in separate blocks. In any one block, one feature could assume one of 9 different levels (durations: 55 - 95 ms in 5 ms steps; pitch mean F0: 0.7 - 1.4 kHz logarithmically spaced; ILD: -8 to +8 dB at an average binaural level of 64 dB, in 2 dB steps; consonant: /da/-/ba/, blended at ratios, from

0.1:0.9 to 0.9:0.1, in equidistant steps; see Table 1), and the remaining features fixed to an intermediate level.

Our roving oddball paradigm is adapted from Garrido et al., (2008), but differs in a few details. First, unlike the previous study, our paradigm used two speech syllables (/ba/ and/da/) instead of the pure tones. Second, we used a different number of deviant step sizes. The previous study used 7 frequency steps, from 500 to 800 Hz in 50 Hz intervals, while we used 9 different levels for each condition. Finally, there were a few differences in stimulus parameters. Garrido et al used stimulus durations of 70 ms and inter-stimulus intervals (ISIs) of 500 ms, with around 200 trials presented for each experimental run. In our study, the stimulus duration was 55-95 ms, the ISI 300 ms, and the deviant and standard presentation yielded around 100 trials for each condition. In our roving oddball paradigm, CV stimuli were presented in continuous trains. The first CV stimulus of a train constitutes an auditory deviant, which becomes a standard after a few repetitions in a train (see Figure 2). After a number of repeat presentations drawn randomly from a uniform interval of 3 to 40 repeats, a new CV stimulus is chosen at random from the 9 possible stimulus levels shown in Table 1. The first occurrence of this new stimulus thus constitutes a "deviant" from the last stimulus of the previous train, which can be of greater or lesser magnitude depending on how much the last standard and the next deviant differ in level. This new stimulus then starts a new train of randomly chosen length, during which it becomes the new "standard". In the analysis, neural responses were compared between the first (deviant) and last (standard) repetition of each train. This ensures that deviant and standard responses in our dataset have precisely the same stimulus parameters and number of trials. Data were acquired in 2 blocks for each condition. Each block comprised 3800 stimuli, and yielded around 50 standard and deviant presentations for each deviance magnitude. Therefore, a total of around 100 standards and deviant AEPs were averaged

per condition in the analysis. Blocks for different conditions were run in a pseudo-random order, and to avoid potential deviant condition order effects, blocks were presented in a different order in each rat.

Table 1. The stimulus features and levels used in the experiment

Level	Duration (ms)	Consonant	ILD (contra- vs. ipsilateral)	Pitch (F0, kHz)
Low	55	/da/ 10% - /ba/ 90%	-8 dB	0.7071
	60	/da/ 20% - /ba/ 80%	-6 dB	0.7711
	65	/da/ 30% - /ba/ 70%	-4 dB	0.8409
	70	/da/ 40% - /ba/ 60%	-2 dB	0.9170
Intermediate	75	/da/ 50% - /ba/ 50%	0 dB	1
	80	/da/ 60% - /ba/ 40%	+2 dB	1.0905
	85	/da/ 70% - /ba/ 30%	+4 dB	1.1892
	90	/da/ 80% - /ba/ 20%	+6 dB	1.2968
High	95	/da/ 90% - /ba/ 10%	+8 dB	1.4142

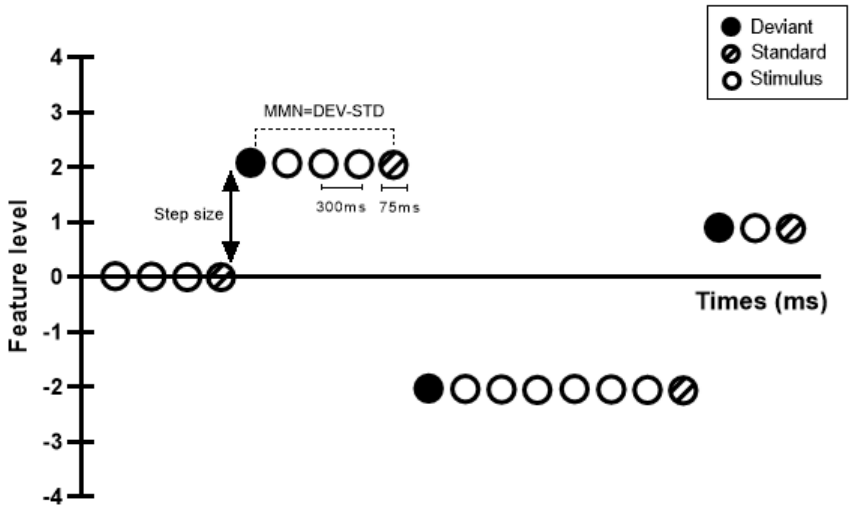

Figure 2. Schematic representation of the stimulation sequences.

CV syllable sounds, represented by circles, are presented in trains of 3 to 40 repeat presentations. At the end of each train, stimulation continues with a new train starting at a different stimulus level. The first stimulus in each train (solid circles) is therefore a deviant sound, while the last (hatched circles) represents a standard sound, after the brain has adapted to the new train. In this example, the first deviant differs from the previous sequence by two levels. The second deviant differs by four levels.

2.3 ECoG data acquisition and pre-processing

EcoG recordings were performed under anaesthesia. The length of each condition was about 20 minutes. ECoG signals were recorded at a sampling rate of 24,414 Hz using a Viventi EcoG electrode array (Woods et al., 2018) connected to a Tucker Davis Technologies (TDT) PZ5 neurodigitizer and RZ2 real-time processor controlled by BrainWare. Stimuli were presented using a TDT RZ6 multiprocessor at a rate of 48,828 Hz. To extract ERPs, the recorded electrode signals were first low pass filtered at a cutoff frequency of 45 Hz using a 5^{th} order Butterworth filter, and downsampled to 1,000 Hz. The pre-processed signals were then epoched by extracting 300 ms long voltage traces from -50 ms to +250 ms relative to the onset of each syllable in the sequence.

The epoched traces were baseline corrected by subtraction of the mean voltage observed over the 50 ms before sound onset, and linearly detrended (Salisbury, 2012). To remove outliers, we calculated a standard deviation of the voltage fluctuation in each trial (SD_i) and rejected trials with SD_i beyond the median ± 3 SD of all SD_i values. This resulted in rejecting 1.08% ± 0.12% of trials (mean ± SD across rats).

Due to individual variations in anatomy and electrode placement, we could not assume that the signals recorded from a particular electrode channel in two different animals would necessarily reflect the activity of precisely matched, equivalent neuronal populations in each rat (see also a quantitative analysis below). We therefore calculated principal spatial and temporal components of the measured responses, and used them to reduce the dimensionality of the data. We then performed statistical analyses on these reduced data, which integrate response patterns over time or space, and are therefore less sensitive to positional mismatches than single electrode and time point data (Lan et al., 2010). Specifically, per rat and acoustic feature, we calculated single-trial difference waveforms between deviant and standard stimuli by subtracting the ECoG amplitude at each channel and time point corresponding to a standard stimulus from the amplitude corresponding to a deviant stimulus. The single-trial difference waveforms were then averaged across trials, resulting in a channel-by-time matrix of amplitude difference values. This matrix was entered into a principal component analysis, resulting in orthogonal temporal and spatial components, ordered from highest to lowest by the amount of variance they explain in the original data. The temporal and spatial components were analyzed separately. For each rat and condition, we took the first N components describing in total at least 99% variance, and calculated their weighted sum (where each component is weighted by the amount of explained variance). These weighted sums of temporal and spatial components are shown in Figure 3.

To test whether the time-series of neural responses to acoustic deviants differ from those to standards, we applied the corresponding weighted summed spatial component to the original data (channel × time × trial), obtaining amplitude values for each time point and trial that effectively summarized all ECoG channels according to how strongly they reflect the difference waveform. This analysis was performed separately for each rat and each condition. First, the resulting single-trial data were averaged across all trials and entered per time point into a repeated-measures ANOVA with one random factor (rat) and two fixed factors (mismatch: deviant vs. standard; condition: duration, consonant, ILD, pitch). Tests were corrected for multiple comparisons across all 300 time points (from -50 to 250 ms relative to tone onset) at a false discovery rate of 0.05 (Benjamini and Hochberg, 1995). Second, having established differences between responses to standards and deviants across all conditions (see Results), we tested whether mismatch responses to different acoustic features have different latencies, by entering the rat-specific latencies (in milliseconds relative to syllable onset) of the peak mismatch responses into a repeated-measures ANOVA with one random factor (rat) and one fixed factor (condition).

Finally, we also tested whether the magnitude of mismatch responses is modulated by deviance magnitude (i.e., by how many steps each deviant differed from the preceding standard) and sign (i.e., by whether each deviant was formed by going up or down a specific acoustic feature). To this end, per time point, we averaged the single-trial difference waveforms (calculated by subtracting the amplitude of responses to standards from those to deviants) separately for each condition, for large vs. small deviance magnitudes, and for different directions of the deviance, resulting in four difference waveforms per rat and acoustic feature. These average mismatch response waveforms were entered into a repeated-measures ANOVA with one random factor (rat) and three fixed

factors (condition; step size: large vs. small; step sign: up vs. down), testing for main and interaction effects. Tests were corrected for multiple comparisons across time points as above.

Beyond comparing the time courses of responses, we also compared the spatial topographies of responses. First, we focused on the spatial principal components and tested for the relative consistency of mismatch response topographies between different acoustic features and rats. To this end, for each rat and condition, we normalized (*z*-scored) each spatial topography and calculated its Euclidean distance (Guggenmos et al., 2018) to either (1) the average spatial topography of the remaining conditions (acoustic features) in the same rat; or (2) the average spatial topography of the remaining rats in the same condition. Second, we tested for the relative consistency of syllable-evoked response (rather than mismatch response) topographies between different acoustic features and rats. Thus, we extracted spatial components of evoked responses (averaged per rat and condition across all trials) in an identical principal component analysis procedure as described above for the mismatch responses. Having obtained the summed weighted spatial components describing the topographies of evoked responses for each rat and condition, we calculated the same two Euclidean distance metrics as for mismatch responses above, this time quantifying the dissimilarity of evoked (rather than mismatch) responses (3) across conditions and (4) across rats. Finally, we quantified the dissimilarity between the evoked and mismatch responses by calculating, per rat and condition, (5) the Euclidean distance between the topography of an evoked response and the topography of a mismatch response. These topography distance measures were compared using paired rank-sum Wilcoxon tests in pair-wise comparisons, correcting for multiple comparisons using a Bonferroni correction.

Finally, we tested whether the topographies of mismatch responses along different acoustic features show consistent spatial distributions across channels. Per rat and condition, we extracted the 2D (dorsal-ventral and rostral-caudal) coordinates of peak electrodes showing (1) the strongest amplitude of the stimulus-evoked response and (2) the strongest amplitude of the mismatch (deviant vs. standard) waveform. To test whether mismatch responses can be mapped at a consistent angle (on the dorsal-ventral and posterior-anterior axes) relative to the spatial peak of the evoked response, we calculated the angle of the unit-normalized vector drawn between the peak of the evoked response and peak of the mismatch response, for each rat and condition. We then tested, per condition, for the clustering of rat-specific angles using the nonparametric Rayleigh z test under the null hypothesis that angles are uniformly distributed on a circle.

3. Results

3.1 Time course and topographic distribution of mismatch responses

Overall, acoustic stimuli evoked differential responses to standards vs. deviants (i.e., mismatch responses) along each of the four acoustic features (Figure 3). Given that the 61 channel ECoG electrode array covered primary and higher-order auditory cortex regions (Doron et al., 2002; Lamas et al., 2017), we could characterize the spatial distribution of the evoked responses (averaged across standards and deviance) as well as the difference between standard and deviant responses in different conditions. The respective spatial topographies are presented in Figure 3B and C. The time courses of mismatch responses, summarizing the principal components explaining > 99% of the observed data, are presented in Figure 4A separately for each condition (acoustic feature).

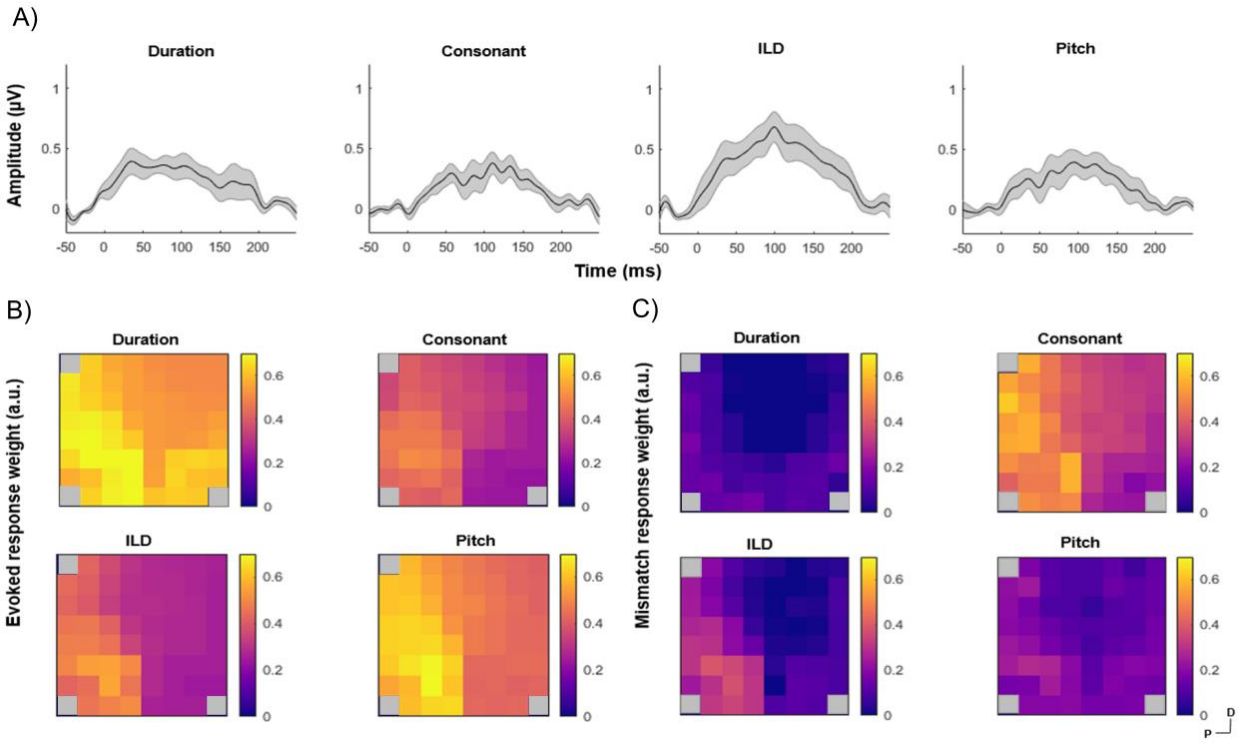

Figure 3. Time courses and topographic distributions of mismatch responses to different acoustic features.

A) The time course of mismatch responses, computed as a weighted sum of principal components of difference waveforms (response to deviant minus response to standard) which explain > 99% of the original variance. The gray-shaded areas denote the SEM across rats. B) The mean topography of the evoked response, averaged across rats. C) The mean topography of the mismatch response, averaged across rats. Gray box indicated reference channel (8,57 and 64). The y-axis represented dorsal to ventral (D-V) and x-axis showed posterior to anterior (P-A).

3.2 Differences in time courses of standard and deviant responses

The spatial principal components of mismatch responses (see above) were applied to the original data (channel × time × trial, separately for each rat and condition) as a form of data dimensionality reduction. The resulting amplitude values, obtained for each time point and trial, effectively summarized all ECoG channels according to how strong they reflected the difference waveform. In a repeated-measures ANOVA, we observed a significant main effect of mismatch between 17

and 193 ms (after correcting for multiple comparisons across time points at a $p < 0.05$, all $F_{1,24} > 7.256$, all uncorrected $p = 0.029$). No significant main effect of condition or interaction between mismatch and condition were observed after correcting for multiple comparisons across time points. Responses to deviants and standards for all four acoustic features are presented in Figure 4A. The latency of peak mismatch responses was not significantly different between conditions ($F_{3,24} = 0.34$, $p = 0.793$).

3.3 No effects of deviance magnitude and direction

Having established significant differences between deviants and standards across conditions, we tested whether the mismatch response depends on step size (the level of deviance: large vs. small) and sign (direction of deviance along with each stimulus feature) in another repeated-measures ANOVA. This analysis revealed no significant main or interaction effects of step size or sign, after correcting for multiple comparisons across time points ($p > 0.05$). Figure 4B shows the mismatch time courses for each condition, step size, and step sign.

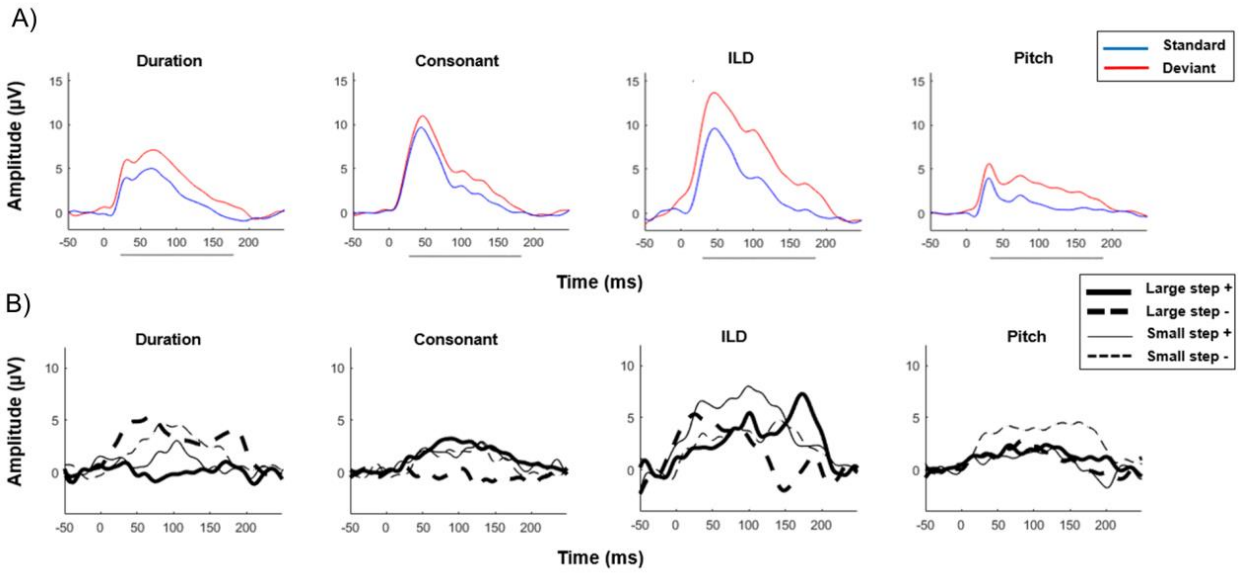

Figure 4. Temporal distributions of mismatch responses through the auditory cortex in different conditions

A) The average waveform of neural responses to acoustic standards (blue) and deviants (red) in each condition (acoustic feature). The horizontal bars (gray) indicate the time points (17ms to 193ms) for which there was a statistically significant main effect of mismatch (deviant vs. standard stimuli; FDR-corrected p-value < 0.05). Since no main or interaction effects of condition were found, these time indices are the same across conditions. B) The average mismatch waveform for each deviant step size (large: thick lines; small: thin lines) and sign (up: solid lines; down: dashed lines). No main or interaction effects were significant after correcting for multiple comparisons across time points.

3.4 Similarity between spatial topographies of mismatch and evoked responses

To compare the topographies between mismatch and evoked responses along different acoustic features, we quantified the (dis)similarity of the spatial distribution of mismatch and evoked responses between conditions (but within rats), between rats (but within conditions), and between each other (within rats and conditions). This analysis (Figure 5A, B) showed that, overall, there was more variability in response distribution between rats than between conditions (evoked responses: $Z = 4.147$, $p < 0.001$; mismatch responses: $Z = 2.938$, $p = 0.003$). Interestingly, while

the variability of mismatch responses between rats was similar to the variability of evoked responses between rats ($p > 0.05$), the variability of mismatch responses between conditions was higher than the variability of evoked responses between conditions ($Z = 3.016$, $p = 0.002$). Furthermore, evoked responses were more dissimilar from mismatch responses (within rats and conditions) than from other evoked responses in different conditions (within rats; $Z = 3.927$, $p < 0.001$). Thus, the spatial distribution of mismatch responses was significantly different from the spatial distribution of the evoked responses.

To test whether the spatial distribution of mismatch responses was systematic across rats, we extracted 2D peak coordinates (dorsal-ventral and anterior-posterior) of the topographies of mismatch responses (deviant vs. standard) and evoked responses (averaged across deviants and standards) per rat and condition. Figure 5B shows the individual rats' peak coordinates of the mismatch response, relative to the peak coordinates of the evoked response, drawn as vectors with a specific length and angle. While mismatch response peaks did not overlap with the evoked response peaks (average ± SEM vector length in mm: duration = 1.157 ± 0.257 mm; consonant = 1.389 ± 0.192 mm; ILD = 1.112 ± 0.396 mm; pitch = 1.298 ± 0.237 mm), the spatial peak of the mismatch response did not have a consistent localization (vector angle) relative to the spatial peak of the evoked response (Rayleigh z test under the null hypothesis of a uniform angle distribution: all $Z < 1.246$). In other words, for individual rats, the strongest mismatch responses for particular features appeared to occur in regions that were "off to one side" from the maximal overall evoked response, which could be interpreted to indicate particular parts of cortex being particularly involved in mismatch detection for that particular sound feature. However, such trends were not consistent across animals, and our data therefore do not support the hypothesis that particular

cortical fields consistently specialize for mismatch detection in specific acoustic features such as pitch, duration or location.

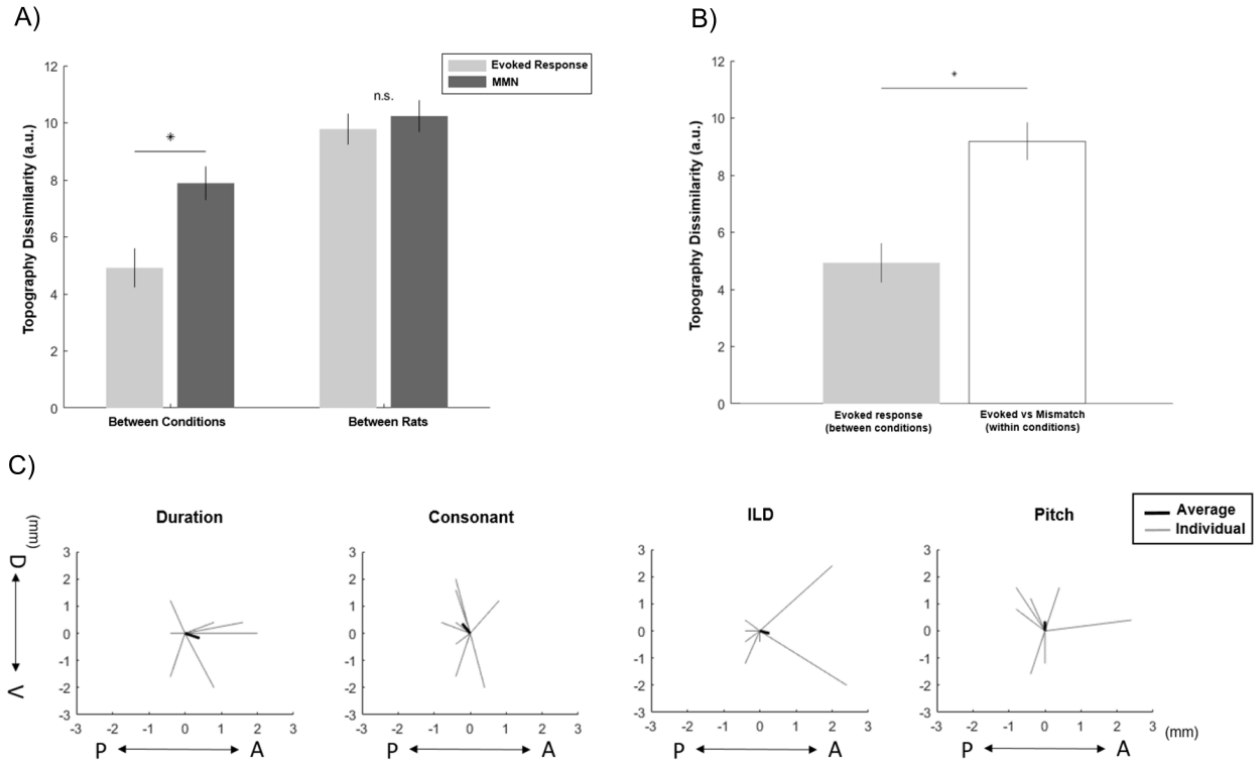

Figure 5. Comparison in spatial neural response between mismatch and evoked responses in different conditions.

A) Average topography dissimilarity between conditions (left bars) and rats (right bars) for the evoked responses (light gray) and mismatch responses (dark gray). B) Average topography dissimilarity evoked responses between conditions (light gray) and evoked responses vs. mismatch responses within conditions (white) C) The difference in peak coordinates between spatial topographies of syllable-evoked responses and mismatch responses shown per condition as a vector for each rat (gray lines) along with the average vector across all rats (black line). The y-axis represented dorsal to ventral (D-V) and x-axis showed posterior to anterior (P-A).

4. Discussion

In this study, we investigated whether violations of different acoustic features (duration, consonant, ILD, and pitch) elicit mismatch responses in anesthetized rats, as well as whether these differential responses to sensory deviants vs. standards can be mapped onto different regions in the auditory cortex. To this end, we recorded neural activity using large-scale surface recordings (ECoG) over primary and higher-order auditory cortex in a roving oddball paradigm where each deviant stimulus was physically identical with a standard. This approach allowed us to find evidence for a differential time course of neural responses to sensory standards and deviants, and quantify the spatial variability of these responses between acoustic features and individual animals.

Traditionally, research on mismatch responses utilizes the classical oddball paradigm, in which a standard stimulus is assigned a high probability (e.g., 90%), and a deviant stimulus is assigned a low probability (e.g., 10%) (Näätänen, 1995; Jaramillo et al., 2000; Garrido et al., 2009b). To control for physical differences between deviant and standard sounds, previous studies introduced several control conditions including a classic-reverse sequence, local-global sequence, and many standards sequences (Bekinschtein et al., 2009; Shiramatsu et al., 2013; El Karoui et al., 2015; Parras et al., 2017). However, they typically necessitate more trials, define standards and oddballs in terms of a global probability rather than local transition probability, and make the effects of stimulus repetition more challenging to study. Therefore, in this study, we used the roving paradigm which controls for physical differences between standards and deviants in an efficient way. The roving oddball paradigm has been used in several clinical studies (Boly et al., 2011; Moran et al., 2013; Uhrig et al., 2014; Rosch et al., 2019) and non-clinical human studies (Garrido et al., 2008; Garrido et al., 2009b). Only a few studies (Komatsu et al., 2015; Takaura and Fujii,

2016) used the roving oddball paradigm to investigate mismatch responses in non-human primates. Although one previous study in guinea pigs (Christianson et al., 2014) characterised mismatch responses in a roving paradigm, this study provides the first demonstration of mismatch responses in rodents in a roving oddball paradigm using multiple acoustic features. In the present study, we found mismatch responses in all acoustic features tested, with no significant latency differences between conditions. Previous studies in rodent models using simple acoustic stimuli (pure tones) observed frequency and duration deviants (Ruusuvirta et al., 1998; Astikainen et al., 2006; Tikhonravov et al., 2008; 2010; Astikainen et al., 2011). Compared to these reports, our data show earlier onset times and longer latency ranges of mismatch responses. These discrepancies may be accounted for by differences in oddball sequences and stimulus types. Our results are consistent with previous studies using syllable stimuli (Ahmed et al., 2011; Komatsu et al., 2015), which revealed earlier onset times of the mismatch response than previous studies using simple acoustic stimuli. This suggests that processing more complex sounds results influence the latency of mismatch responses in the rat model. However, previous human studies showed not only overall later-onset times and longer latency ranges of mismatch responses (Jaramillo et al., 2000; Grimm et al., 2004; Tardif et al., 2006), but also differences in latencies between stimulus features (Takegata et al., 1999). This may be explained by mismatch detection eliciting higher-order cognitive processes such as stimulus anticipation in humans (Näätänen et al., 2007), which may be more difficult to observe under anesthesia or in rodent models more generally. This suggests that the modulation of mismatch responses by stimulus features might depend on the complexity of brain networks.

While the overall differences between neural responses evoked by deviants and standards were robust for all stimulus features tested in this study, we did not observe further modulations of

mismatch response amplitude by deviance magnitude or direction. The results of previous studies which tested for differences in deviance direction or magnitude (Shelley et al., 1991; Jaramillo et al., 2000; Umbricht et al., 2005; Takegata et al., 2008; Shiramatsu et al., 2013) were mutually inconsistent, which may have resulted from differences in step size between stimuli. For instance, duration step sizes have ranged from 50 ms (Shelley et al., 1991; Umbricht et al., 2005) to 1 ms (Jaramillo et al., 2000), or from 50% (Umbricht et al., 2005) to 16% (Roger et al., 2009b). Generally, rodent studies have found significant differences in mismatch response amplitude for large vs. small magnitude of deviance when step sizes were above 33% (Roger et al., 2009b; Lipponen et al., 2019), while in humans these modulations can be observed even for 10% step sizes(Jaramillo et al., 2000). While our selection of step sizes was based on previous literature (Walker et al., 2009; Ahmed et al., 2011; Parras et al., 2017), taken together, the results of this and previous studies suggest that future experiments in the anesthetized rodent models should include relatively large deviance magnitudes.

On the other hand, we observed differences in spatial distribution between the mismatch responses and evoked responses, as well as between mismatch responses to different acoustic features. The variability of evoked responses between rats was similar to the variability of mismatch responses between rats, suggesting that overall the spatial distribution of mismatch responses and evoked responses could be identified with similar reliability. However, mismatch responses were more variable between conditions (acoustic features) than evoked responses. While a previous study in the ferret auditory cortex (Bizley et al., 2009) demonstrated that multiple perceptual attributes of sound, including the pitch, timbre and spatial location of the sound, can be represented by overlapping populations of neurons distributed among multiple auditory fields, our results suggest that mismatch responses to different acoustic feature are even more heterogeneous than evoked

responses. This is broadly consistent with a previous microelectrode study in rat (Shiramatsu et al., 2013), reporting that the focal activation loci of evoked response were in the auditory core region, while mismatch responses had broader spatial distributions, including both core and belt regions. However, our study also shows that mismatch responses vary considerably across conditions and individual rats, consistent with a previous study in humans (Phillips et al., 2015) which found that, while multiple regions in the frontotemporal network are sensitive to auditory mismatch, the overall activity levels in single temporal and frontal regions do not differentiate between mismatch responses to different acoustic features. Taken together, these results support the notion that mismatch responses have a broad and diverse spatial distribution, similar to other higher-order functions (Herbert et al., 1991; Malmierca, 2003).

In our study, we aimed to record signals from primary and higher-order auditory regions in order to map any regional differences between mismatch responses to multiple acoustic features. However, as can be appreciated in Figure 3B and C, there was little evidence for a functional separation between regions based on the topography of the sound-evoked response, which made it difficult to delineate primary from higher-order regions based on the activity evoked by the stimuli used in this study. Even more importantly, as reported in Figure 5C, there was a large degree of heterogeneity in the relative topographies of evoked and mismatch responses across rats, again supporting the idea that the topography of observed (evoked or mismatch) responses would have been a poor ground for functional segregation of areas into primary and higher-order cortices. Based on these results, we decided not to perform separate analyses on subgroups of channels. However, it remains a possibility that other functional mapping methods (e.g. based on tonotopic-mapping) would help in delineating the functional neuroanatomy of auditory regions in terms of primary and higher-order cortices, as indeed suggested by previous work (Higgins et al., 2010;

Shiramatsu et al., 2013; Nieto-Diego and Malmierca, 2016). Therefore, it will be important to address functional separation between primary and higher-order regions in future work.

In conclusion, in the present study, we observed mismatch responses to several acoustic features, with different spatial patterns between evoked and mismatch responses, as well as between mismatch responses to different acoustic features. However, the spatial distributions of mismatch responses showed considerable differences between individual rats, suggesting that while prediction errors to different stimulus features are mediated by different neural populations (Auksztulewicz et al., 2018a; Stefanics et al., 2019), these populations may not be consistently grouped into separate auditory fields at the level of primary and higher-order auditory cortices. Future work should combine neural recordings of mismatch responses in awake animals engaged in a change detection task, as well as compare the results of rodent and human studies in an identical paradigm. This would yield insights into the functional role of mismatch response variability, and shed light on the extent to which the predictive processing of different perceptual dimensions of sounds is preserved in evolution.

Acknowledgments

Funding for this project was provided by the Hong Kong General Research Fund (# 11100518) and a grant from European Community / Hong Kong Research Grants Council Joint Research Scheme. The ECoG electrodes used in these experiments were kindly provided by Prof Jonathan Viventi of Duke University. We would like to thank Ms Luo Dan and Ms Kongyan Li for their help with the recording experiments.

Chapter 3.

Do auditory mismatch responses differ between acoustic features?

HyunJung An[1], Shing Ho Kei[1], Ryszard Auksztulewicz[1,2]*, Jan W. Schnupp[1]*

1 Department of Neuroscience, City University of Hong Kong, Hong Kong

2 Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

* Equal contribution

* Correspondence:
Prof. Jan W.Schnupp.
xxxxxxxx@cityu.edu.hk

Dr. Ryszard Auksztulewicz
rxxxxxx@gmail.com

Keywords: Electroencephalography (EEG), Mismatch negativity (MMN), Predictive coding, Auditory processing, Multivariate decoding

Abstract

Mismatch negativity (MMN) is the electroencephalographic (EEG) waveform obtained by subtracting event-related potential (ERP) responses evoked by unexpected deviant stimuli from responses evoked by expected standard stimuli. While the MMN is thought to reflect an unexpected change in an ongoing, predictable stimulus, it is unknown whether MMN responses evoked by changes in different stimulus features have different magnitudes, latencies, and topographies. The present study aimed to investigate whether MMN responses differ depending on whether sudden stimulus change occur in pitch, duration, location or vowel identity respectively.

To calculate ERPs to standard and deviant stimuli, EEG signals were recorded in normal-hearing participants (N=20; 13 males, 7 females) who listened to roving oddball sequences of artificial syllables. In the roving paradigm, any given stimulus is repeated several times to form a standard, and then suddenly replaced with a deviant stimulus which differs from the standard. Here, deviants differed from preceding standards along one of four features (pitch, duration, vowel or interaural level difference). The feature levels were individually chosen to match behavioral discrimination performance.

We identified neural activity evoked by unexpected violations along all four acoustic dimensions. Evoked responses to deviant stimuli increased in amplitude relative to the responses to standard stimuli. A univariate (channel-by-channel) analysis yielded no significant differences between MMN responses following violations of different features. However, in a multivariate analysis (pooling information from multiple EEG channels), acoustic features could be decoded from the topography of mismatch responses, although at later latencies than those typical for MMN. These

results support the notion that deviant feature detection may be subserved by a different process than general mismatch detection.

1. Introduction

Neural activity is typically suppressed in response to expected stimuli and enhanced following novel stimuli (Carbajal and Malmierca, 2018). This effect is often summarized as a mismatch response, calculated by subtracting the neural response waveform to unexpected deviant stimuli from the response to expected standard stimuli. Auditory deviance detection has been associated with a human auditory-evoked potential, the mismatch negativity, occurring at about 150–250 ms from sound change onset (Näätänen, 2007; Garrido et al., 2008). The principal neural sources of the MMN are thought to be superior temporal regions adjacent to the primary auditory cortex, as well as frontoparietal areas (Doeller et al., 2003; Chennu et al., 2013). Initially, the MMN was interpreted as a correlate of pre-attentive encoding of physical features between standard and deviant sounds (Doeller et al., 2003). However, more recent studies have led to substantial revisions of this hypothesis, and currently, the most widely accepted explanation of the MMN is that it reflects a prediction error response.

An important theoretical question remains whether mismatch signaling has a domain-general or domain-specific (feature-dependent) implementation in the auditory processing pathway. A recent study using invasive recordings from the cortical surface (Auksztulewicz et al., 2018b) demonstrated that neural mechanisms of predictions regarding stimulus contents ("what") and timing ("when") can be dissociated in terms of their topographies and latencies throughout the frontotemporal network, and that activity in auditory regions is sensitive to interactions between different kinds of predictions. Additionally, biophysical modeling of the measured signals has shown that predictions of contents and timing are best explained either by short-term plasticity or

by classical neuromodulation respectively, suggesting separable mechanisms for signaling different kinds of predictions. However, these dissociations might be specific to predictions of contents vs. timing, which may have fundamentally different roles in processing stimulus sequences (Friston and Buzsaki, 2016).

Interestingly, an earlier magnetoencephalography (MEG) study (Phillips et al., 2015) provided evidence for a hierarchical model, whereby violations of sensory predictions regarding different stimulus contents were associated with similar response magnitudes in auditory cortex, but different connectivity patterns at hierarchically higher levels of the frontotemporal network. This result is consistent with the classical predictive coding hypothesis in which reciprocal feedforward and feedback connections at the lower levels of the hierarchy are thought to signal prediction errors and predictions regarding simple sensory features, but hierarchically higher levels are thought to signal more complex predictions and prediction errors, integrating over multiple features (Kiebel et al., 2008). Several studies, however, reported independent processing of prediction violations along different acoustic features or sound dimensions. An earlier study (Giard et al., 1995) investigated the neural correlates of mismatch processing across three different acoustic features (frequency, intensity, and duration). Mismatch responses to each feature were source-localised by fitting equivalent current dipoles to EEG signals, and the results indicated that violations of different features can be linked to dissociable sources, suggesting the involvement different underlying populations. Similar conclusions have been reached in another set of studies (Schröger, 1995; Paavilainen et al., 2001), which quantified the additivity of MMN to changes along different acoustic features, either in isolation or by combining two or more features. In these studies, the MMN response to violating two features could largely be reproduced by adding the MMN responses to violating two single features, suggesting that the latter are mutually independent. A

more recent study has combined these two approaches (source localization and additivity analyses), demonstrating partial independence of three different timbre dimensions (Caclin et al., 2006). The notion that mismatch responses to violations of different features are mediated by independent mechanisms is also supported by studies showing that MMN (as well as the later P3a component) typically decreases following two identical deviants presented in direct succession, but remains stable following two deviants which vary from the standard along different features (for a review, see Rosburg et al., 2018).

However, in most previous studies (Giard et al., 1995; Schröger, 1995; Paavilainen et al., 2001; Phillips et al., 2015; Rosburg et al., 2018), physical differences between deviants and standards were not behaviorally matched across different features or participants, raising the possibility that differences in mismatch-evoked activity might to some extent be explained by differences in stimulus salience (Shiramatsu and Takahashi, 2018). This was also the case in the more recent studies on MMN responses to multiple acoustic features (Phillips et al., 2015) or in previous roving paradigms (Garrido et al., 2008). Interestingly, a recent study investigating the MMN to acoustic violations along multiple independent features in the auditory cortex of anaesthetised rats (An et al., 2020) revealed that the topography of MMN signals was highly diverse across not only acoustic features but also individual animals, even though several sources of inter-subject variability (e.g. electrode placement) were better controlled than in typical non-invasive studies, suggesting that the spatial resolution of non-invasive methods such as EEG or MEG might not be sufficient for mapping more subtle differences between mismatch responses to violations of different features. The few EEG studies that did use behaviourally matched deviant sounds across different features either used very small sample sizes (N=8 (Deouell and Bentin, 1998)) or were limited to relatively specialised perceptual characteristics (e.g. different timbre features: Caclin et al., 2006). In contrast,

our study used a larger sample size (N=20) and manipulated relatively general sound dimensions (location, pitch, duration, and syllable identity). Our primary goal was to test whether mismatch responses to violations of different features differ in magnitude or latency, in an attempt to replicate previous studies (Deouell and Bentin, 1998). However, in addition to testing the effects of acoustic feature on the MMN time-course in a mass-univariate analysis (i.e., on an electrode-by-electrode basis), we also aimed at decoding acoustic features from differences in MMN topography in a multivariate analysis (i.e., pooling signals from multiple electrodes).

2. Materials and Methods

2.1 Participants

Twenty volunteers (13 males and 7 females; mean age 23.9 years old) enrolled in the study upon written informed consent. All participants self-reported as having normal hearing and no history of neurological disorders, and all but two were right-handed. All participants but one were native Hong Kong residents, and their mother tongue was Cantonese. A musical training questionnaire indicated that 16 participants had no musical training, and the remaining participants had less than four years' experience in playing a musical instrument. Participants were seated in a sound-attenuated and electrically shielded room in front of a computer screen. They were instructed to fixate on a fixation cross displayed on the screen during the acoustic stimulation. All experimental procedures were approved by the Human Subjects Ethics Sub-Committee of the City University of Hong Kong.

2.2 Stimuli

The present study employed a roving oddball paradigm in which auditory deviants could differ from preceding standards along one of four independent acoustic features. Specifically, we manipulated two consonant-vowel (CV) syllable stimuli, /ta/ and /ti/ (Retsa et al., 2018), along the following independent acoustic features: duration, pitch, interaural level difference (ILD) or vowel (An et al., 2020). Prior to the EEG recording, per participant, we estimated the feature interval yielding ~80% behavioral performance by employing a 1-up-3-down staircase procedure. In each staircase trial, two out of three stimuli, chosen at random, were presented at a mean level of a given feature (e.g., a 50/50 vowel mixture or a 0 dB ILD) while the third stimulus was higher or lower than the mean level by a certain interval. Participants had to indicate which stimulus was the "odd one out". Following three consecutive hits, the interval decreased by 15%; following a mistake, the interval increased by 15%. Each participant performed 30 staircase trials for each feature (Figure 1.(B)). For the roving oddball stimulus sequences, the stimulus duration was set to 120 ms and the inter-stimulus intervals (ISIs) were fixed at 500 ms. Stimuli formed a roving oddball sequence: after 4-35 repetitions of a given stimulus (forming a standard), it was replaced with another (deviant) stimulus, randomly drawn from the set of 5 possible levels (Figure1. (A)). Roving oddball sequences corresponding to different features were administered in separate blocks, in a randomized order across participants. The total number of stimuli in each block was approximately 2000, including 200 deviant stimuli and 200 corresponding (immediately preceding) standards.

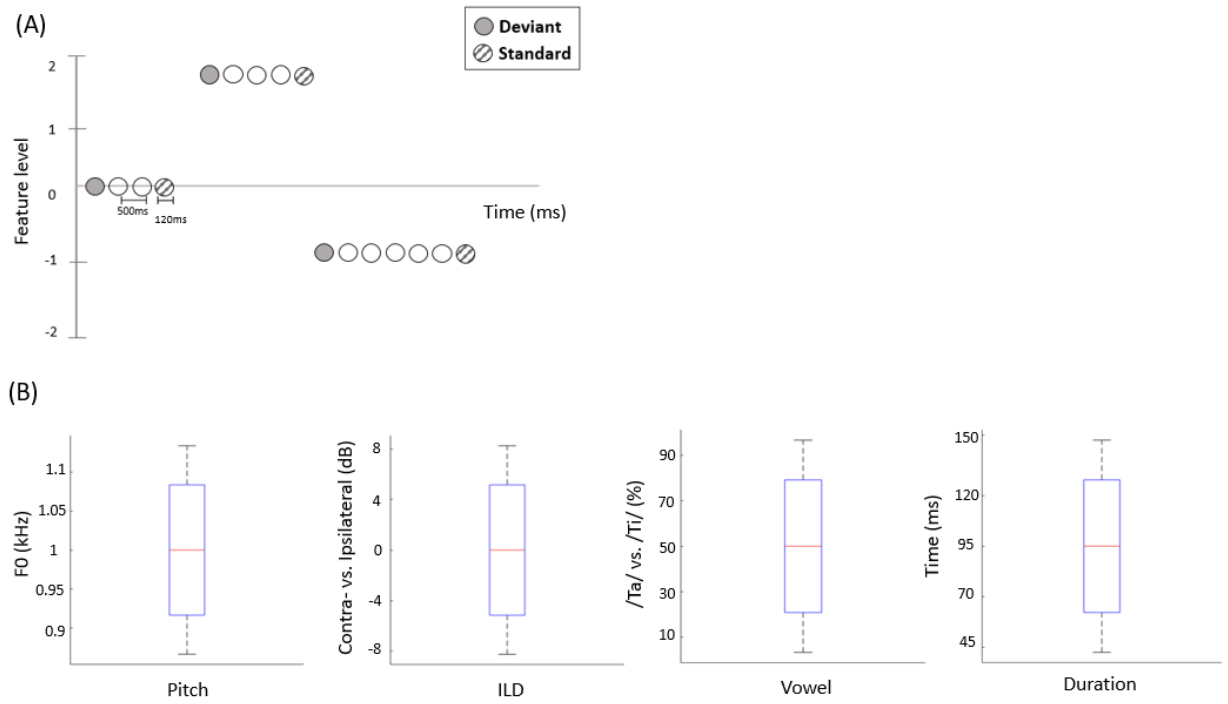

Figure 1. (A) Schematic representation of the stimulation sequences. The first stimulus in each train (solid circles) represents a deviant sound, while the last (hatched circles) represents a standard sound. (B) The range of each acoustic feature used to construct stimuli in the EEG experiment. Red line indicates the median value of each feature (across participants), blue bars and black whiskers represent mean and SD of upper and lower ranges across participants.

2.3 Experimental procedure

We recorded signals from 64 EEG channels in a 10-20 system using an ANT Neuro EEG Sports amplifier. EEG channels were grounded at the nasion and referenced to the Cpz electrode. Participants were seated in a quiet room and fitted with Brainwavz B100 earphones, which delivered the audio stimuli via a MOTU Ultralite MK3 USB soundcard at 44.1 kHz. EEG signals were pre-processed using the SPM12 Toolbox for MATLAB. The continuous signals were first notch-filtered between 48 and 52 Hz and band-pass filtered between 0.1 and 90 Hz (both filters: 5th order zero-phase Butterworth), and then downsampled to 300 Hz. Eye blinks were automatically detected using the Fp1 channel, and the corresponding artefacts were removed by subtracting the two principal spatiotemporal components associated with each eye blink from all

EEG channels (Ille et al., 2002). Then, data were re-referenced to the average of all channels, segmented into epochs ranging from -100 ms before to 400 ms after each stimulus onset, baseline-corrected to the average pre-stimulus voltage, and averaged across trials to obtain ERPs for deviants and standards for each of the four acoustic features.

2.4 Data analyses

First, to establish the presence of the MMN response, we converted the EEG time-series into 3D images (2D spatial topography × 1D time-course) and entered them into a general linear model (GLM) with two factors (random effect of mismatch: deviant vs. standard; fixed effect of participant), corresponding to a paired t-test. Statistical parametric maps were thresholded at an uncorrected $p < 0.005$, and the resulting spatiotemporal clusters of main effects were tested for statistical significance at the family-wise error corrected threshold $pFWE < 0.05$, taking into account the spatiotemporal correlations and multiple comparisons across channels and time points.

In an additional control analysis, we have tested whether the mismatch responses observed in this study were modulated by adaptation effects, which have been shown to be especially prominent in the N1 range (Baldeweg et al., 2004). To this end, per standard stimulus (i.e., the last stimulus in a sequence of identical stimuli), we have calculated the number of stimuli separating it from the preceding deviant (i.e., the first stimulus in a sequence of identical stimuli). If our results were indeed confounded by adaptation, the difference between responses evoked by deviants vs. standards should be modulated by the number of stimuli preceding each deviant. To test this hypothesis, we have regressed out the number of preceding stimuli from single-trial standard-evoked responses (using two regressors: a linear regressor, coding for the actual number of preceding stimuli, and a log-transformed regressor, approximating empirically observed

adaptation effects; e.g. (Baldeweg et al., 2004), and subjected the residuals to the remaining univariate analysis steps (i.e., averaging the single-trial responses to obtain ERPs, and performing statistical inference while correcting for multiple comparisons across channels and time points).

Then, to test whether MMN amplitudes differed between stimulus features, ERP data were entered into a flexible-factorial GLM with one random factor (participant) and two fixed factors (mismatch: deviant vs. standard; feature: pitch, duration, ILD, and vowel), corresponding to a repeated-measures 2 × 4 ANOVA. Statistical significance thresholds were set as above.

Finally, to test whether mismatch responses can be used to decode the violated acoustic features, we subjected the data to a multivariate analysis. Prior to decoding, we calculated single-trial mismatch response signals by subtracting the EEG signal evoked by each standard from the signal evoked by the subsequent deviant. Data dimensionality was reduced using PCA (principal component analysis), resulting in spatial principal components (describing channel topographies) and temporal principal components (describing voltage time-series), sorted by the ratio of explained variance. Only those top components which, taken together, explained 95 % of the original variance, were retained for further analysis. In decoding acoustic features, we adopted a sliding window approach, integrating over the relative voltage changes within a 100 ms window around each time-point (Wolff et al., 2020). To this end, per channel and trial, the time segments within 100 ms of each analysed time-point were down-sampled by binning the data over 10 ms bins, resulting in a vector of 10 average voltage values per component. Next, the data were de-meaned by removing the component-specific average voltage over the entire 100 ms time window from each component and time bin. These steps ensured that the multivariate analysis approach was optimised for decoding transient activation patterns (voltage fluctuations around a zero mean)

at the expense of more stationary neural processes (overall differences in mean voltage) (Wolff et al., 2020).

The binned single-trial mismatch fluctuations were then concatenated across components for subsequent leave-one-out cross-validation decoding. Per trial and time point, we calculated the Mahalanobis distance (De Maesschalck et al., 2000) (scaled by the noise covariance matrix of all components) between the vector of concatenated component fluctuations of this trial (test trial) and four other vectors, obtained from the remaining trials, and corresponding to the concatenated component fluctuations averaged across trials, separately for each of the four features. The resulting Mahalanobis distance values were averaged across trials, separately for each acoustic feature, resulting in 4 × 4 distance matrices. These distance matrices were summarized per time point and participant as a single decoding estimate, by subtracting the mean off-diagonal from diagonal terms (Figure. 3(A)).

In a final analysis, since we have observed univariate mismatch responses as well as multivariate mismatch-based feature decoding at similar latencies (see Results), we have tested whether these two effects are related. To this end, we performed a correlation analysis between single-trial decoding estimates (i.e., the relative Mahalanobis distance values between EEG topography corresponding to mismatch responses following violations of the same vs. different features), and single-trial MMN amplitudes. We calculated Pearson correlation coefficients across single trials, per channel, time point, and participants. The resulting correlation coefficients were subject to statistical inference using statistical parametric mapping (one-sample t-test; significance thresholds as in the other univariate analysis, corrected for multiple comparisons across time points and channels using family-wise error).

3. Results

Taken together, in this study, we tested whether auditory mismatch responses are modulated by violations of independent acoustic features. First, consistent with previous literature (Doeller et al., 2003; Garrido et al., 2008), we observed overall differences between the ERPs evoked by deviant stimuli vs. standard stimuli, in a range typical for MMN responses as well as at longer latencies (Figure 2. (A)). Specifically, the univariate ERP analysis confirmed that EEG amplitudes differed significantly between deviants and standards when pooling over all the acoustic features tested. This effect was observed over two clusters: the central EEG channels showed a significant mismatch response between 115 and 182 ms (cluster-level $pFWE < 0.001$, $Tmax = 3.94$), while posterior channels showed a significant mismatch response between 274 and 389 ms (cluster-level $pFWE < 0.001$, $Tmax = 5.46$), within the range of a P3b component. A control analysis, in which we controlled for single-trial adaptation effect to the standard tones, yielded a virtually identical pattern of results as the original analysis (two significant clusters of differences between responses to deviants vs. standards: an earlier cluster between 130 and 143 ms over central channels, cluster-level $pFWE < 0.001$, $Tmax = 15.48$, and a later cluster between 317 and 327 ms over posterior channels, cluster-level $pFWE < 0.001$, $Tmax = 17.48$).

Although the ERP time-courses differed between deviant and standard stimuli when pooling over violations of different acoustic features, a univariate (channel-by-channel) analysis revealed no significant differences in the amplitudes or time-courses of mismatch responses between independent stimulus features (Figure 2. (B)). These results are consistent with a previous study (Phillips et al., 2015) which found that multiple deviant stimulus features (frequency, intensity, location, duration, and silent gap) were not associated with differences in activity in the auditory

regions, but instead were reflected in more distributed activity patterns (frontotemporal connectivity estimates).

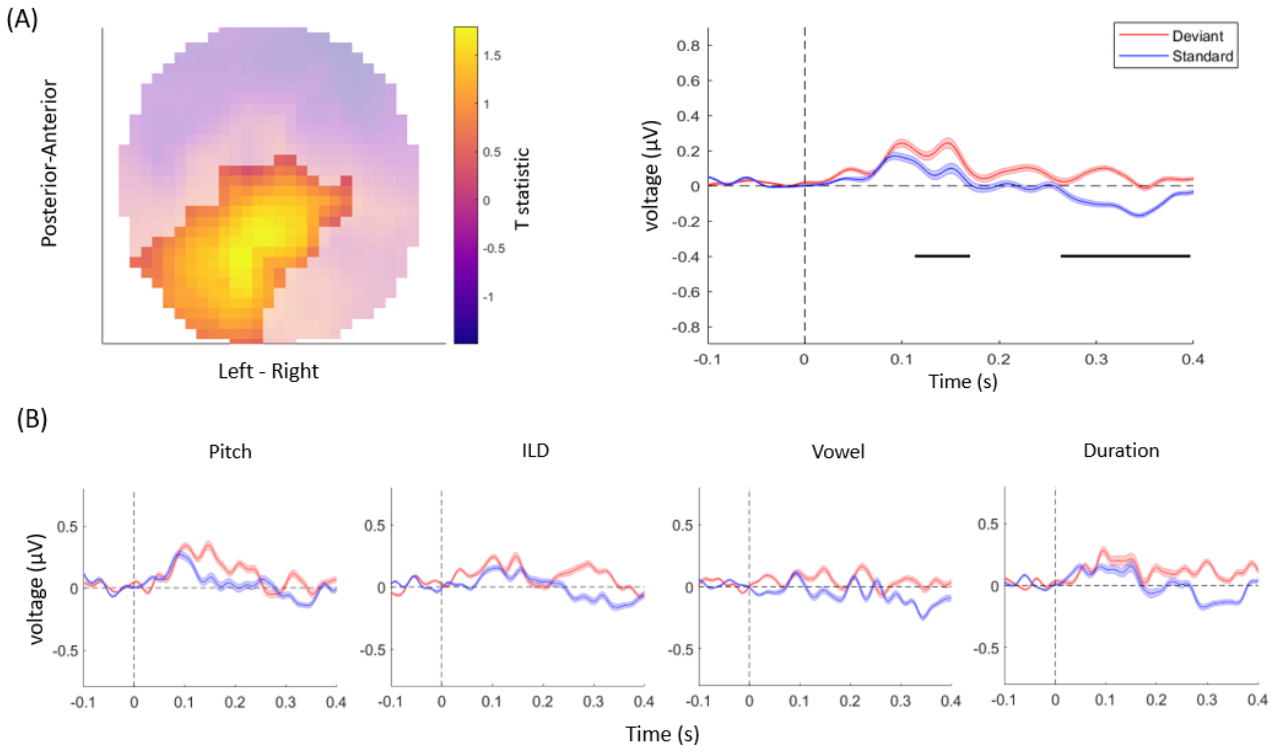

Figure 2. (A) The topography (left) and time-course (right) of the mismatch response. The highlighted topography cluster represents the significant difference between deviants and standards. Based on this cluster, the average waveform of the evoked response is plotted separately for auditory standards (blue) and deviants (red). The horizontal bars (black) indicate time points with a significant difference between deviants and standards. Shaded areas denote SEM (standard error of the mean) across participants. (B) The average response to acoustic standards (blue) and deviants (red) for different feature conditions, extracted from the same cluster as in (A). No interaction effects were significant after correcting for multiple comparisons across channels and time points.

The resulting decoding time-courses of each participant were entered into a GLM and subject to one-sample t-tests, thresholded at an uncorrected $p < 0.05$ and correcting for multiple comparisons across time points at a cluster-level $pFWE < 0.05$. In this analysis, significant acoustic feature decoding was observed between 247 and 350 ms relative to tone onset (cluster-level $pFWE = 0.000$,

Tmax = 2.77). (Figure 3. (B)). Thus, when pooling information from multiple EEG channels, acoustic features could be decoded from the topography of mismatch responses, although at later latencies than typical for MMN.

Since we have observed both univariate mismatch responses and multivariate mismatch-based feature decoding at late latencies (univariate: 274 - 389 ms; multivariate: 247 – 350 ms), we have performed an additional single-trial correlation analysis to test whether these two effects are related. This analysis (Figure 3. (C)) has yielded no significant clusters of correlation coefficients between single-trial mismatch amplitudes and decoding estimates, while correcting for multiple comparisons across channels and time points (Tmax = 3.74, all pFWE > 0.005).

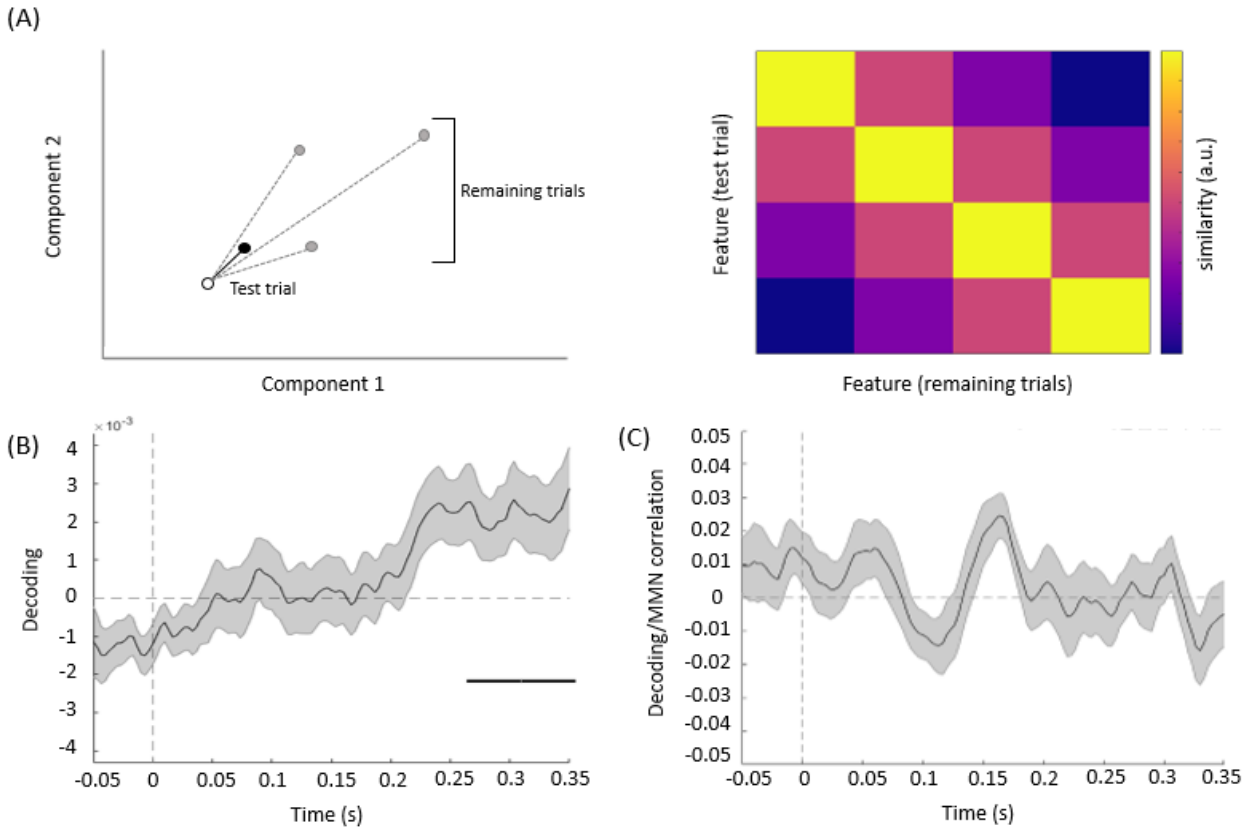

Figure 3. (A) Decoding methods. Left panel: for each trial, we calculated the Mahalanobis distance, based on multiple EEG components (here shown schematically for two components), between the mismatch response in a given (test) trial (empty circle) and the average mismatch responses based

on the remaining trials (black circle: same feature as test trial; grey circles: different features). Right panel: after averaging the distance values across all trials, we obtained 4 by 4 similarity matrices between all features, such that high average Mahalanobis distance corresponded to low similarity between features. Based on these matrices, we summarized feature decoding as the difference between the diagonal and off-diagonal terms. (B) Multivariate analysis. The average time course of the decoding of acoustic features based on single-trial mismatch response. The gray-shaded area denotes the SEM across participants, and the horizontal bar (black) shows the significant time window. (C) Decoding vs. MMN correlation analysis. Plot shows the time-series of mean correlation coefficients between single-trial decoding estimates and single-trial MMN amplitudes, calculated for Cz/Cpz channels and averaged across participants (shaded areas: SEM across participants). No significant correlations were observed when correcting for multiple comparisons across channels and time points.

4. Discussion

In this study, since a univariate analysis of interactions between mismatch signals and acoustic features might not be sensitive enough to reveal subtle and distributed amplitude differences between conditions, we adopted a multivariate analysis aiming at decoding the violated acoustic feature from single-trial mismatch response topographies. This demonstrated that acoustic features could be decoded from the topography of mismatch responses, although at later latencies than typical for MMN (Figure 3. (B)). An earlier oddball study (Leung et al., 2012) examined ERP differences to violations of four features (frequency, duration, intensity, and interaural difference). The study found that frequency deviants were associated with a significant amplitude change in the middle latency range. This result indicated that deviant feature detection may be subserved by a different process than general mismatch detection. Consistent with this notion, another study has used magnetoencephalography to identify mid-latency effects of local prediction violations of simple stimulus features, and contrasted them with later effects of global prediction violations of stimulus patterns (Recasens et al., 2014). Taken together, these studies would suggest that, in paradigms where multiple acoustic features vary independently (such as here), a plausible pattern

of results would be that independent feature predictions should be mismatched at relatively early latencies, since an integrated representation is not required. Here, however, we found feature-specificity in the late latency range, rather than in the mid-latency range. The discrepancy between our results and the previous studies might be explained by different stimulus types. While the previous studies used simple acoustic stimuli, here we used complex syllable stimuli, possibly tapping into the later latencies of language-related mismatch responses, as compared to MMN following violations of non-speech sounds.

Speech sounds have been hypothesized to be processed in separate streams which independently derive semantic information ("what" processing) and sound location ("where" processing) (Kaas and Hackett, 2000; Tian et al., 2001; Schubotz et al., 2003; Camalier et al., 2012; Kusmierek and Rauschecker, 2014). In most animal studies, the hierarchical organization of the auditory cortex has been linked to a functional distribution of stimulus processing, such that core (hierarchically lower) regions respond preferentially to simple stimuli, whereas belt and other downstream (hierarchically higher) regions respond to more complex stimuli such as band-passed noise and speech (Rauschecker et al., 1995; Recanzone et al., 2000; Rauschecker and Tian, 2004; Kusmierek and Rauschecker, 2009; Rauschecker and Scott, 2009). This is supported by evidence functional magnetic resonance imaging (fMRI) studies in humans (Binder et al., 2000) showing that earlier auditory regions (Heschl's gyrus and surrounding fields) respond preferentially to unstructured noise stimuli, while progressively more complex stimuli such as frequency-modulated tones show more lateral response activation patterns. In that study, speech sounds showed most pronounced activations spreading ventrolaterally into the superior temporal sulcus. This result supports a hierarchical model of auditory speech processing in the human auditory cortex based on complexity and integration of temporal and spectral features. Based on this notion, the relatively

long latency of neural responses compared to previous studies using pure tones might be partially explained by the fact that we used spectrally and temporally complex speech stimuli.

However, our results can also be explained in terms of a hierarchical deviance detection system based on predictive coding (Kiebel et al., 2008). On this account, neural responses supporting the lower and higher hierarchical stages communicate continuously through reciprocal pathways. When exposed to repetitive stimuli, the bottom-up (ascending) sensory inputs can be "explained away" by top-down (descending) connections mediating prediction signaling, resulting in weaker prediction error signaling back to the hierarchically higher regions. Substituting the predicted standard with unpredicted deviant results in a failure of top-down suppression by prior predictions. This leads to an increased prediction error signaling back to higher regions, providing an update for subsequent predictions. As a result, the later and more distributed activity patterns might reflect higher-order prediction errors, signalled to regions integrating multiple stimulus features and representing the entire range of stimuli likely to appear in a particular context.

In conclusion, the present study identified functional dissociations between deviance detection and deviance feature detection. First, while mismatch responses were observed at latencies typical for the MMN as well as at longer latencies, channel-by-channel analyses revealed no robust differences between mismatch responses following violations of different acoustic features. However, we demonstrate that acoustic features could be decoded at longer latencies based on fine-grained spatiotemporal patterns of mismatch responses. This finding suggests that deviance feature detection might be mediated by later and more distributed neural responses than deviance detection itself.

5. Author Contributions

HyunJung An: Formal analysis, Writing original draft, Conceptualization, Conducted experiment. Shing Ho Kei: Conducted experiment, Formal analysis. Ryszard Auksztulewicz: Formal analysis, Supervision, Project administration, Conceptualization. Jan W.H. Schnupp: Project administration, Conceptualization, Investigation, Supervision.

6. Funding

This work has been supported by the European Commission's Marie Skłodowska-Curie Global Fellowship (750459 to R.A.), the Hong Kong General Research Fund (11100518 to R.A. and J.S.), and a grant from European Community/Hong Kong Research Grants Council Joint Research Scheme (9051402 to R.A. and J.S.).

Chapter 4.

Neural correlates of auditory pattern learning in the auditory cortex

Kang, H.[1], Auksztulewicz, R.[1,2], An. H. J.[1], Abichacra, N.[1], Sutter, M. L.[3], and Schnupp, J. W. H.[1*]

[1]Department of Neuroscience, City University of Hong Kong

[2]Neuroscience Department, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

[3]Center for Neuroscience, and Section of Neurobiology, Physiology and Behavior, University of California, Davis, California

* Correspondence:
Jan W. H. Schnupp
xxxxxx@cityu.edu.hk

Keywords: auditory perception[1], learning[2], electrocorticography[3], rat[4], auditory cortex[5].

Abstract

Learning of new auditory stimuli often requires repetitive exposure to the stimulus. Fast and implicit learning of sounds presented at random times enables efficient auditory perception. However, it is unclear how such sensory encoding is processed on a neural level. We investigated neural responses that are developed from a passive, repetitive exposure to a specific sound in the auditory cortex of anesthetized rats, using electrocorticography. We presented a series of random sequences that are generated afresh each time, except for a specific reference sequence that remains constant and re-appears at random times across trials. We compared induced activity amplitudes between reference and fresh sequences. Neural responses from both primary and non-primary auditory cortical regions showed significantly decreased induced activity amplitudes for reference sequences compared to fresh sequences, especially in the beta band. This is the first study showing that neural correlates of auditory pattern learning can be evoked even in anesthetized, passive listening animal models.

1 Introduction

Sensory perception requires correctly recognizing incoming sensory stimuli by extracting relevant information from memory. Such memory can be formed by implicit learning of sensory input through repetitive exposure. Fast memory formation by capturing unique features of sensory signals is thus one key factor for efficient sensory perception, which requires active involvement of primary sensory cortices (Harris et al., 1999; Bao et al., 2004; Gavornik and Bear, 2014; Rosenthal et al., 2016).

In hearing, a series of recent studies reported fast and robust learning of abstract sounds, using a novel experimental paradigm that resembles unsupervised implicit learning of newly presented

acoustic stimuli in auditory scenes (Agus et al., 2010; Luo et al., 2013; Andrillon et al., 2015). In this paradigm, participants were simply asked to detect a within-sequence repetition in random noise samples. Unbeknownst to them, one specific noise sample would re-occur occasionally, and even though the subjects were unaware of this, they nevertheless showed fast, selective improvement in processing the frozen "reference" stimulus, which implies rapid and robust memorization of random features of complex sounds. Such behavioral improvement for the re-occurring sound was supported by increased inter-trial coherence of brain responses for the re-occurring stimulus compared to other random stimuli measured by subsequent EEG and MEG studies in humans (Luo et al., 2013; Andrillon et al., 2015). Interestingly, increases in neural coherence could even be observed when the human subjects were in Rapid Eye Movement (REM) or light non-REM sleep during the experiment (Andrillon et al., 2017), suggesting that a neural index related to learning new sounds can be traced even following passive exposure. While these findings provided insights into the neural correlates of implicit learning of new auditory stimuli, further investigations using invasive measurements will be needed to understand the underlying mechanisms. The present study aimed at investigating neural responses shaped by passively presented re-occurring sounds in the auditory cortex using rats as an animal model.

Previous electrophysiological studies have investigated how neurons adapt to re-occurring sounds to understand memory and adaptation processes, by using a simplified experimental paradigm, in which a series of standard sounds (usually pure tones) is disrupted by a presentation of a deviant sound (Garrido et al., 2009a; Malmierca et al., 2014; Nieto-Diego and Malmierca, 2016). Under such paradigms, stimulus specific adaptation (SSA) has been widely reported using comparisons between habituated neural responses to a standard sound against the typically greater responses for a novel, deviant sound. SSA effects have been observed along the auditory pathway, first in

the primary auditory cortex (AC), and then also in non-lemniscal subdivisions of the inferior colliculus (IC) and the medial geniculate body (MGB; Anderson et al., 2009; Ayala and Malmierca, 2013; Parras et al., 2017; Ulanovsky et al., 2003). A more recent study further reported stronger SSA in non-primary AC fields compared to primary AC (Nieto-Diego and Malmierca, 2016). Another study using more complex and realistic sounds has suggested that higher-order regions in the AC, rather than primary fields, may be uniquely susceptible to the adaptation to repeatedly presented realistic auditory inputs (Lu et al., 2018). The study further reported that the adaptation effect was retained after the disruption period from another repetitive presentation of the other sound input in the AC. These results point to the active involvement of the AC in learning and adaptation to ongoing or predictable sounds, which is thought to play a role not only in encoding stimuli, but also their context (Bar-Yosef et al., 2002; Skipper, 2014; Lu et al., 2018), as well as the prefrontal cortex (Casado-Roman et al., 2020). However, while previous studies compared neural responses evoked by occasional deviants relative to consecutively presented standards, such constant presentation of a single sound is not sufficient to fully explain our ability of fast implicit learning for newly presented sounds. Instead, recognizable re-occurring sounds typically appear occasionally, interspersed with other random, non-repeating sounds, and yet listeners learn them without much effort.

In the present study, instead of the classical paradigm of constant representations of a single sound, we adapted an experimental paradigm (Agus et al., 2010) to intermittently present frozen "reference" sequences among other random sequences. The aim of the present experiment was to look for a physiological correlate of the "learning" of the frozen sequence that can occur even during passive exposure in the AC of anesthetized rats, using electrocorticography (ECoG) as a first step for identifying neurophysiological markers. We focused on investigating neural

characteristics that emerged by learning re-occurring auditory patterns across primary and non-primary auditory fields within the AC. We particularly controlled for physical differences between stimuli by comparing pure induced neural responses, non-phase locked responses to the stimulus computed after the time-frequency decomposition rather than evoked neural responses. By doing so, we could minimize any observed effect to be drawn from characteristics of the stimulus itself and focus on the neural modulations induced by higher-order processing (Klimesch et al., 1998; David et al., 2006). Our results show more attenuated induced activity amplitudes for the re-occurring sounds compared to other sounds, in both primary and non-primary fields, especially in the beta frequency band.

2 Methods

2.1 Animal Subjects

Six female adult Wistar rats (age = 8 - 21 weeks, mean = 12.5, std = 4.42, weight = 257 – 315 g) were acquired from the Chinese University of Hong Kong. Experimental procedures were approved by the City University Animal Research Ethics Sub-Committee and conducted under license by the Department of Health of Hong Kong [Ref. No. (19-31) in DH/SHS/8/2/5 Pt.5].

2.2 Stimuli

We generated sequences of acoustic stimuli shown schematically in Fig. 1A. Each sequence consisted of five segments. Sequences could be made up either of 0.24 s long dynamic random chords (DRCs) or of 0.2 s long white noise (WN) snippets. Sequences either consisted of the same segment repeated five times (repeated sequence, RS), or they were non-repeating, random sequences (S). To make it easier to distinguish neural signatures of repetition detection from simple

onset or offset responses, sequences were bracketed with additional “head” and “tail” segments, which were always generated afresh, and ramped on or off linearly. Segments were joined with 5 ms ramping overlaps to avoid transients. Sequences were presented in blocks with inter-sequence-interval of 0.6. One block contained 100 unique RS and S sequences each, as well as one “frozen RS” and one “frozen S” sequence, which were presented 100 times each in each block. Adopting the nomenclature of Agus et al. (2010) we refer to the frozen sequences as “references”, or “RefRS” and “RefS” respectively. Thus one block consisted of a shuffled series of 400 stimuli, with 100 different S and RS sequences and 1 unique RefS and RefRS sequence being presented in random order (see Fig. 1B). For each block, sequences were generated anew with new random seeds.

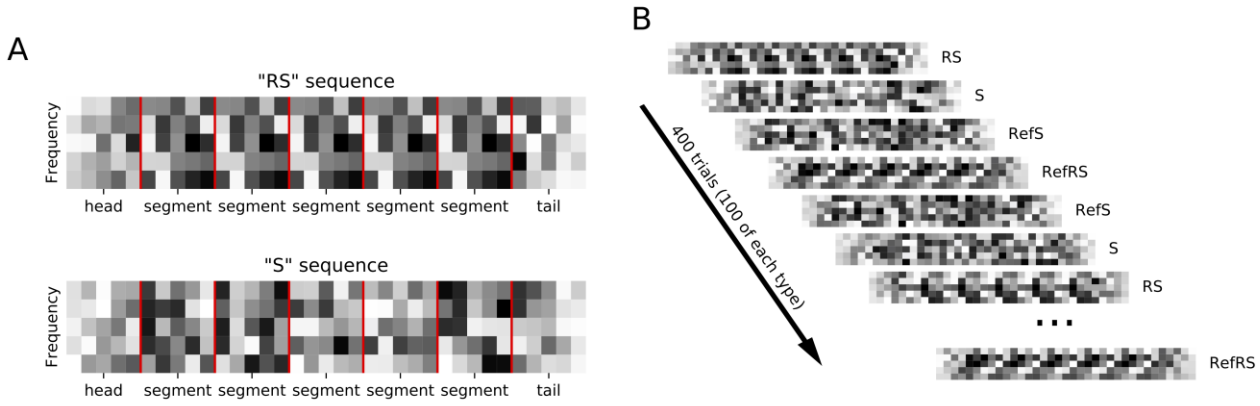

Figure 1. A: Sequences were composed of random spectral pattern (DRC or white noise) segments (marked by red vertical lines) which were either repeated 5 times in a row (RS sequence) or non-repeating (S sequence). Ramped, random “head” and “tail” segments bracketed each sequence. B: sequences were presented in blocks of 400 trials. Each block contained 100 sequences, each of unique R and S sequences as well as repeated “reference” RefRS and RefS sequences, which were presented in random order.

DRC sequences consisted of 12 chords of superimposed 20 ms pure tones at 15 log-spaced frequencies from 500 to 20,000 Hz. The level of each tone was randomly drawn from a uniform 50-90 dB SPL range to have mean 70 dB SPL, generating random spectro-temporal patterns

characteristic of each DRC. The WN sequences consisted of Gaussian noise snippets, generated from different random seed values. As DRC segments comprise more salient spectral contrasts than WN segments, we expected within-sequence repetition to be easier to detect in DRC than in WN sequences.

2.3 Experimental procedure

We recorded responses to five blocks of DRC sequences and five blocks of WN sequences from ECoG arrays placed onto the right auditory cortex (AC). Anesthesia was induced using Ketamine (80 mg/kg) and Xylazine (12 mg/kg, Intraperitoneal injection; i.p.) and maintained with Urethane (20%, 7.5 µl/g, i.p.). Urethane anesthesia minimizes NMDA receptor blockage and closely resembles REM and stage II nREM sleep-like status (Pagliardini et al., 2012). Dexamethasone (0.2 mg/kg, i.p.) was injected to prevent inflammation. Adequate anesthesia was confirmed by regular testing for the suppression of the toe pinch withdrawal reflex. Body temperature was kept at 36 ± 1 °C with a heating pad. The rat was placed in a stereotaxic frame and the head was fixed with hollow ear bars to allow the delivery of auditory stimuli. We measured the auditory brainstem responses (ABRs) in each ear to confirm that the rats had normal hearing sensitivity (click thresholds < 20 dB SPL). The right AC was exposed by a rectangular 5 x 4 mm craniotomy which extended from 2.5 to 7.5 mm posterior from Bregma, with its medial edge 2.5 mm from the midline (Polley et al., 2007). A 61-channel ECoG array (Woods et al., 2018) was connected to a Tucker Davis Technologis (TDT) PZ5 neurodigitizer and RZ2 real-time processor, and placed on the exposed cortex. Sound stimuli were presented via a TDT RZ6 multiprocessor through the hollow ear bars at a sampling rate of 48,828 Hz, and ECoG responses were recorded at 24,414 Hz using BrainWare software.

The correct placement of the ECoG array was confirmed by recording frequency responses to 100 ms pure tones at frequencies (500 – 32k Hz, ¼ octave steps) at 70 dB SPL to obtain frequency tuning curves of individual electrodes and a mapping of the best frequency across at the recording site. Given that the ECoG electrodes are rather large and their spacing is relatively wide relative to the reported dimensions of tonotopic fields of the rat described in the literature, the FRA maps obtained did not show clear tonotopic gradients, but they nevertheless revealed physiological features of a frequency response topography of the AC, which were reproducible from animal to animal. In particular, we were able to verify that tentative A1 areas have distinct frequency gradients from low to high frequencies, while the tentative non-A1 areas (SRAF) have frequency gradients from high to low frequencies (from caudal to rostral; Fig. 2). These findings are consistent with previous studies from other laboratories (e.g., Nieto-Diego and Malmierca, 2016).

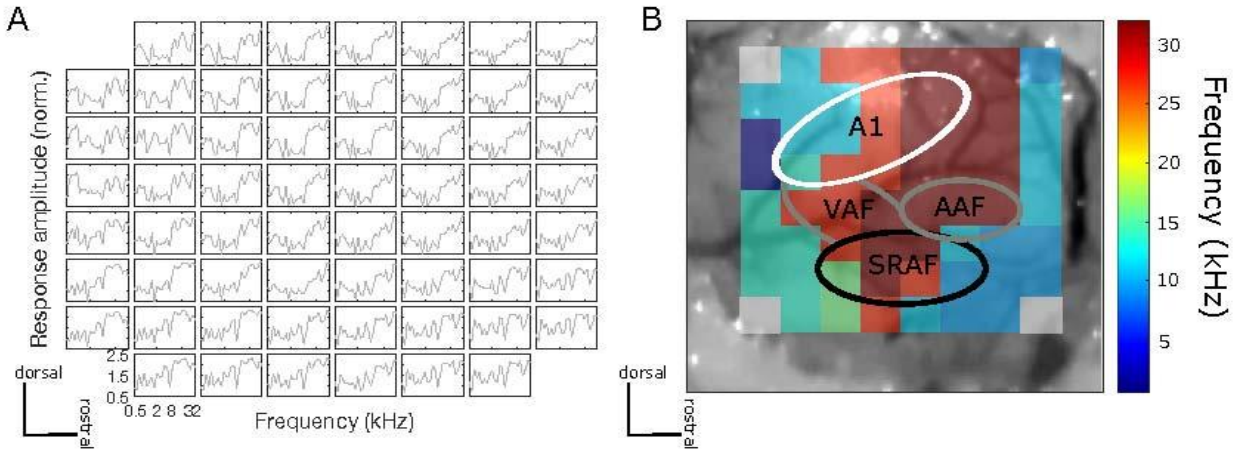

Figure 2. A: An example of frequency tuning curves for each electrode at the recording site to 100 ms pure tone at different frequencies at 70 dB SPL. B: An example central frequency gradient across the recorded AC site. Tonotopic gradients of tentative A1 (low to high characteristic frequency from caudal to rostral) and non-A1 (high to low characteristic frequency from caudal to rostral) areas were observed.

2.4 Data analyses

Acquired neural responses were pre-processed to obtain event-related potentials (ERPs) for each channel and condition for each rat. ERPs were used to look for differences across the four conditions (S, RS, RefS, RefRS), using the time-frequency analysis described below. To calculate ERPs of each channel, ECoG signals were low-pass filtered (2^{nd} order zero-phase Butterworth) at 45 Hz, downsampled to 1,000 Hz, and re-referenced to the common mean. Time points at which signal values exceeded ± 3 std of the mean signal across time were identified as outliers and removed (i.e. replaced by linear interpolation from neighboring points, and detrending as described in (de Cheveigne and Arzounian, 2018). Signals for each trial were then epoched from -100 ms to 1600 ms relative to the onset of each sequence. Epochs for each condition were averaged to compute mean ERPs for each channel. To reduce data dimensionality, as well as minimize the effect of individual variations in electrode placement between rats, we subjected each rat's channel-by-time ERP matrix (averaged across conditions) to a principal component analysis (PCA) and ordered components from the highest to the lowest amount of variance. We selected the top components (in order of variance explained) describing at least 99% variance and calculated the weighted sum of the spatial components to quantify the evoked response topography with reduced variabilities across rats. A visual inspection of regional response differences per rat from the obtained topography revealed that channels with the lowest response weights were mainly around A1 areas while channels with the highest response weights were mainly around non-A1 areas. Thus, we grouped top response-weighted channels as a tentative non-A1 cluster, and the bottom response-weighted channels as a tentative A1 cluster for further analysis of regional differences. Since the number of channels included in each group did not affect the result, we grouped the channels into the top 30 channels for the non-A1 cluster and the rest for the A1 cluster.

Next, to characterize the differences in induced responses to reference sequences (RefRS and RefS) compared to fresh sequences (RS and S), we ran a time-frequency analysis of single-trial ECoG signals using Morlet wavelets implemented in the FieldTrip toolbox for Matlab (frequency range: 4 - 80 Hz in 2 Hz steps; 400 ms fixed time window; Billig et al., 2019) for each rat. The time-frequency power spectrum of each trial was rescaled by subtracting the time-frequency spectrum of the average ERP for the same condition (i.e., evoked power) on a logarithmic scale. This subtraction yielded an estimate of induced activity amplitude, whereby the responses in each individual trial did not have to be precisely time-locked to the stimulus (Hartmann et al., 2012). Therefore, the induced response differs from the ERP by focusing on the oscillation of spectral power rather than on phase-locked responses to the stimuli. The resulting single-trial induced responses were log-scaled and averaged across trials. After obtaining average time-frequency power spectra for each rat, channel group, and stimulus condition, we ran two cluster-based permutation paired *t*-tests (as implemented in FieldTrip) on RefRS vs. RS stimuli and on RefS vs. S across rats as independent observations, with 1000 iterations per test. The results of this statistical analysis would then suggest that the observed effects would be largely consistent across rats.

3 Results

First, a cluster-based permutation paired *t*-test with 1,000 iterations revealed no significant differences of ERP amplitudes (averaged across channels) between RefRS and RS conditions or RefS and S conditions, either for DRC or for WN sequences.

For both DRC and WN stimuli, channels presumed to be primary auditory cortex (A1) showed lower evoked response weights (averaged across all trials and conditions) than channels presumed to be non-primary auditory cortex, mainly from around suprarhinal auditory field (SRAF; Fig. 3A).

Based on the evoked response weights, we grouped channels into two clusters, A1 and non-A1 clusters, based on the evoked response weights for further analyses on comparing induced activity amplitudes differences across conditions. A time-frequency analysis of induced activity revealed robust differences between pairs of Ref and non-Ref conditions for both clusters for DRC, but not for WN.

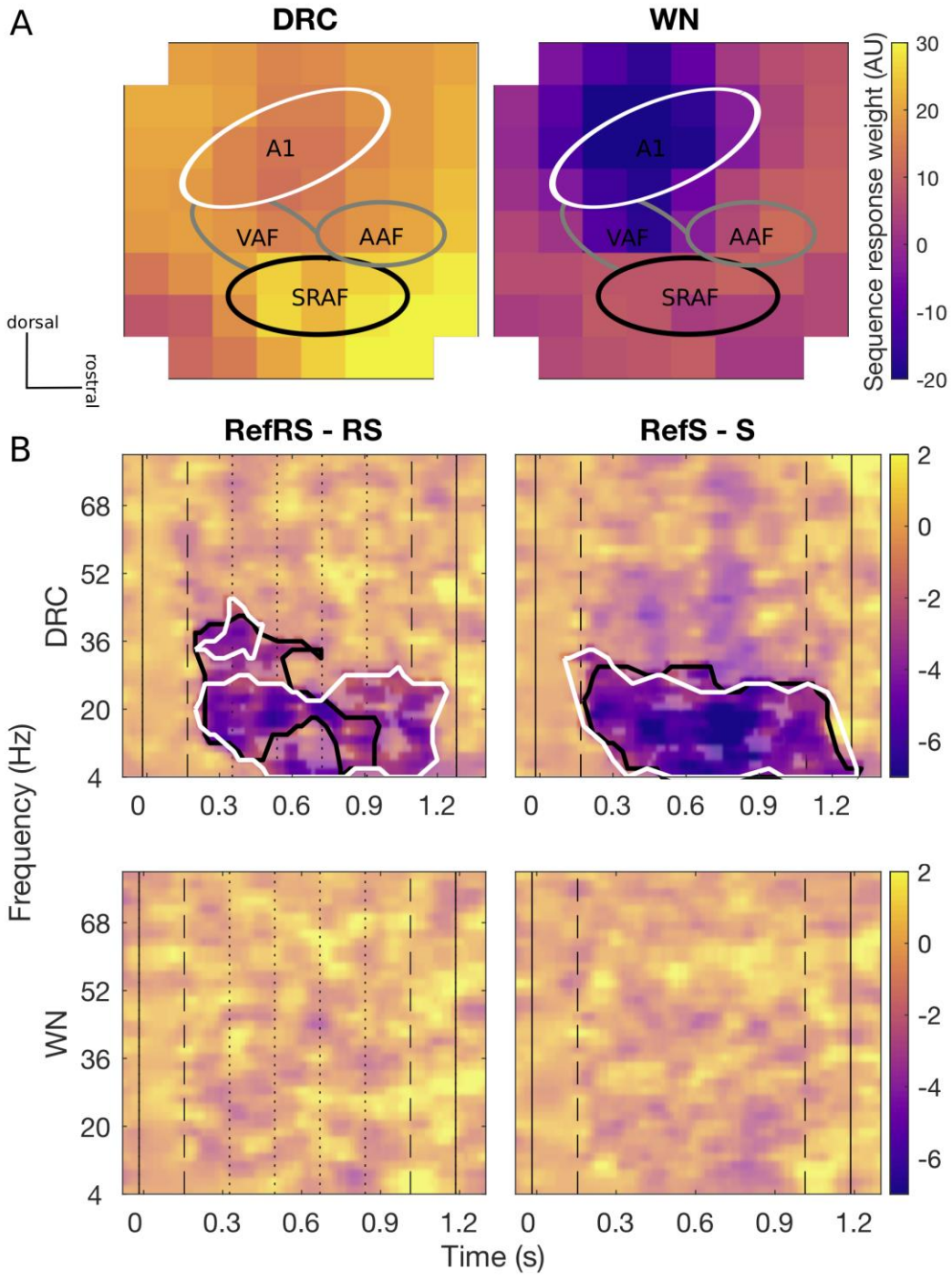

Figure 3. A: Spatial topography maps (methods of Nieto-Diego and Malmierca, 2016 and Polley et al., 2007) of average evoked responses collected by 8 x 8 ECoG for DRC (left) and WN (right) sequences respectively. Each pixel represents individual ECoG channel of a 8 x 8 grid placed over the AC area. Colorscale is fixed for both sound types. Tentative subfields of the AC are marked with relevant labels (A1, VAF, AAF, and SRAF). The main A1 cluster (white line) shows slightly

lower evoked response weights, and the putative SRAF cluster (black line) showed generally greater evoked response weights. Overall evoked responses to DRC sequences were stronger than for WN. In both cases, greater evoked responses were generally found in the non-A1 clusters. B: Differences in average time-frequency induced power spectra for RefRS minus RS (left) and RefS minus S (right) pairs for DRC (top) and WN (bottom). Black solid vertical lines indicate sound onset and offset, dashed vertical lines indicate reference sequence onset and offset, and dotted vertical lines on the left panel indicate within-sequence segment boundaries. Black contours in DRC spectra indicate time-frequency areas where a significant difference between conditions was observed for the non-A1 cluster, and white contours indicate the areas with a significant difference observed for the A1 cluster (cluster-based $p < 0.05$). No significant power difference was observed for both pairs in WN.

We first focused on neural activity in the non-A1 cluster induced by DRC stimuli, based on the hypothesis that perceptual learning of complex stimuli may primarily modulate activity in higher-order regions. When comparing time-frequency responses between RefRS and RS conditions, we observed significantly decreased power for RefRS vs. RS during the sequence presentation, emerging from the onset of RefRS mostly in the beta band (10 – 40 Hz; $T_{min} = -19.41$, $T_{max} = -2.57$, all cluster-based p's < 0.05). This was especially pronounced during the first three segments of the sequences (Fig. 2B). In the RefS vs. S comparison, decreased power for RefS was also observed from the RefS onset to the sound offset across the theta, alpha, and beta band (4 – 30 Hz; $T_{min} = -8.54$, $T_{max} = -2.58$, all cluster-based p's < 0.05). In the A1 cluster, power decrease for RefRS vs. RS was observed in a similar frequency range (4 – 30 Hz; $T_{min} = -19.82$, $T_{max} = -2.58$, all cluster based p's < 0.05) to the non-A1 cluster, but persisted for a longer time period (from sequence onset to sound offset), mostly in the beta band. Power decrease for RefS vs. S was observed for similar time period and frequency bands to the non-A1 cluster (4 – 30 Hz; $T_{min} = -9.98$, $T_{max} = -2.57$, all cluster-based p's < 0.05). For WN sequences, we did not observe any significant differences in time-frequency response spectra between either RefRS vs. RS or RefS vs. S comparisons.

4 Discussion

We assessed distinct neural correlates of implicit learning processes through repetitive passive exposure to a specific auditory sequence. We compared neural dynamics of re-occurring sequences with the same acoustic characteristics (RefRS and RefS) and a group of other sequences that were presented only once (RS and S), by computing induced activity amplitude of neural signals recorded from primary (A1) and higher-order auditory cortex. We observed decreased induced activity amplitude throughout the stimulus sequence for RefRS and RefS compared to RS and S, mainly in the beta band, both for A1 and non-A1 channel clusters, but only for DRC stimulus sequences which contain more salient acoustical features compared to WN. This finding suggests an active involvement of both primary and non-primary AC in the implicit learning of complex auditory patterns.

Unlike most previous studies that computed differences between evoked responses as an index of learning (Lim et al., 2016; Lu et al., 2018), we did not observe any significant difference in ERPs across conditions. This result, however, was expected in our study as Ref sequences were presented in a passive listening setting with a complex and unpredictable experimental design. Previous neuroimaging studies in humans under similar paradigms also mainly focused on comparing inter-trial coherence rather than ERP differences between RefRS and RS or RefS and S as similar ERPs for RefRS and RS were observed in most cases (Andrillon et al., 2015; 2017; Luo et al., 2013). Thus, we focused on comparing induced response power obtained from each trial after the time-frequency decomposition. We found distinct response patterns for both of re-occurring sequences (RefRS and RefS) from the beginning of the sequence presentation when compared to fresh sequences (RS and S) for DRC. Each test block contained different, randomly generated re-

occurring sequences, thus there was no build-up effect along successive blocks. Such difference suggests that within-sequence repetitions are not a requirement for the learning process, as long as the sequence contains salient information to be learned. Both effects in our study were observed mostly in the beta band, which has been implicated in sensory memory (Haenschel et al., 2000; Scholz et al., 2017) and sensory predictions (Pearce et al., 2010; Auksztulewicz and Friston, 2016; Auksztulewicz et al., 2017).

The effect of attenuated induced activity amplitude for RefRS over RS and RefS over S was found from both A1 and non-A1 channel clusters. The effect mostly overlapped between the two clusters, especially for RefS. Interestingly, for RefRS, significant power differences in the non-A1 cluster started to diminish already after the first three segment repetitions – unlike the power difference in the A1 cluster which lasted towards the end of the sequence (Fig. 2B). Although further investigation is required, we hypothesize that repeated segments within RS may also have been learned, resulting in no distinctive difference between RefRS and RS to be observed towards the end of the sequence. It could be possibly due to increased suppression to re-occurring segments in putative non-lemniscal (non-A1) clusters relative to lemniscal (A1) clusters (Parras et al., 2017). could further indicate that acoustic features presented in re-occurring brief segments that are as short as 200 ms can be effectively learned and recognized. Such characteristic was only observed in the non-A1 cluster, suggesting a hierarchical structure of the auditory cortex and indicating a role of higher-order regions in repetition suppression and prediction in a shorter time frame (Auksztulewicz and Friston, 2016; Lim et al., 2016; Lu et al., 2018). The present finding provides further insights into neural responses mediating RefRS learning within a short timeframe. Whether such characteristics remain in a longer-term memory can also be further studied.

One caveat of the present study is that we did not observe a gradual development of the effect across trials, which is one of key factors that confirms the effect as an outcome of learning. Since the data were recorded with ECoG as a first step to verify whether any difference emerges in the AC, the present study focused on the signals accumulated over multiple trials. Further investigation with single and multi-unit recordings would be beneficial to study changes of neuronal activities on a trial by trial level, by separating units which are selective to the Ref sequences from non-selective ones to increase the signal to noise ratio (e.g., Lu et al., 2018).

Finally, although a previous study using the similar paradigm showed neural correlates of RefRS using WN as stimuli (Luo et al., 2013; Andrillon et al., 2015), we did not observe any distinctive characteristic of RefRS and RefS for WN. The repeating segment length of WN in the present study (200 ms) was shorter than previous studies that used WN (500 ms; e.g., Agus et al., 2010). This could be due to a greater saliency that DRC could generate for its recognition compared to WN, especially when such short repeating segment was presented. Other factors such as the segment duration or seamless presentation could also have affected the saliency of the Ref sequence (Agus et al., 2010; Andrillon et al., 2017; Kang et al., 2017), which requires further investigation to study features that make certain sequences more 'memorable' than others.

The present experiment was conducted under anesthesia. Anesthetics could affect certain aspects of auditory processing such as spike timing, population activity or frequency tuning, depending on the type of anesthetics (Zurita et al., 1994; Gaese and Ostwald, 2001; Huetz et al., 2009; Noda and Takahashi, 2015). However, a large amount of previous physiology research investigating neural adaptation for re-occurring sounds or information processing has been conducted on animals under anesthesia (e.g., Anderson et al., 2009; Bao et al., 2004), especially using Urethane (e.g.,

Astikainen et al., 2011; Astikainen et al., 2014; Lipponent et al., 2019). These studies, as well as our results, suggest that neural responses under anesthesia carry important information that are highly correlated with sensory perception. In the present study, the usage of Urethane was chosen to minimize any effect from anesthetics (Capsius and Leppelsack, 1996; Hara and Harris, 2002; Curto et al., 2009). Furthermore, our findings suggesting distinct neural traces for Ref sequences over the AC in rats under anesthesia are comparable to the findings observed from human neuroimaging study during REM sleep (Andrillon et al., 2017). Thus, the present findings provide important initial findings on neural correlates during such passive, implicit learning in AC. Certain discrepancies between the present findings and previous human studies (e.g., no significant difference found from WN presentations) could be further studied by conducting experiments in awake animals.

In summary, the present study showed distinctive neural traits for re-occurring abstract auditory patterns that provide salient acoustic features (DRC) in the auditory cortex. While decreased induced activity amplitudes in the beta band observed throughout the auditory cortex suggests that both A1 and non-A1 areas are involved in encoding the information of re-occurring acoustic stimuli, such memory encoding in non-A1 areas could be processed in a shorter time frame. In this study we report, for the first time, a neural correlate of this type of memory formation in an easy to use, passive listening animal model, which should greatly facilitate further investigation into underlying neural mechanisms.

5 Author Contributions

HK, RA, NA, MS, and JS designed the study and ran pilots, HK and HA conducted the experiments, HK and RA analyzed the data, HK wrote the first draft of the manuscript, all authors revised the manuscript.

6 Funding

This work has been supported by Fyssen Foundation (study grant to HK), the Hong Kong General Research Fund (11100518 to RA and JS), a grant from European Community/Hong Kong Research Grants Council Joint Research Scheme (9051402 to RA and JS), and European Commission's Marie Skłodowska-Curie Global Fellowship (750459 to RA), the National Institutes of Health (DC002514 to MLS), and the Fulbright Scholar Program (to MLS).

Chapter 5

Summary & Conclusions

The major findings of my thesis can be summarized as follows:

1. Mismatch responses to prediction violations along different acoustic features are more heterogeneous than sound-evoked responses.

   In auditory processing, MMN reflects the neural signal resulting from a comparison between previous sensory experience and a novel auditory input. Although previous studies have reported differences in MMN depending on the acoustic features of the deviant (Ahmed et al., 2011; Mahmoudzadeh et al., 2017; Lipponen et al., 2019), they did not compare the mismatch response between regions. This study investigates whether violations of different acoustic features can be mapped onto different regions in the auditory cortex.

   a) This study found that mismatch responses were more spatially distributed across conditions than evoked responses;
   b) While mismatch and evoked responses showed a similar heterogeneity of topographies across animals, mismatch responses were more heterogeneous across acoustic features than evoked responses.

   Thus, the mismatch responses to different acoustic features are likely subserved by different but idiosyncratically distributed neural populations.

2. Deviance feature detection mechanism is dissociable from general mismatch detection.

Building upon the previous study, we tested whether the MMN response shows feature specificity in humans. Most previous studies using different acoustic features were not behaviorally matched across features and were limited to relatively specialized stimulus characteristics (Giard et al., 1995; Schröger, 1995; Paavilainen et al., 2001; Caclin et al., 2006; Phillips et al., 2015; Rosburg et al., 2018). This study used behaviorally matched deviant stimuli across different features to decode acoustic features from the MMN topography in a multivariate analysis.

a) This study showed that, in a univariate analysis, EEG amplitude significantly differed between deviant and standard stimuli (in an earlier cluster over central channels and later cluster over posterior channels) but not between acoustic features.

b) In a multivariate analysis, decoding of acoustic features was observed at later latencies than typical for MMN.

Thus, there is a functional dissociation between deviance detection and deviance feature detection.

3. Neural correlates of auditory learning can be developed from the passive condition. Auditory memory can be formed by implicit learning due to repeated exposure to auditory stimuli. Previous studies have attempted to explain how neurons adapt to repetitive sounds and form a memory representation (Garrido et al., 2009a; Malmierca et al., 2014; Nieto-Diego and Malmierca, 2016). However, the studies compared neural responses evoked by very simple sounds forming unrealistically simple auditory scenes.

This study focused on investigating neural correlates of learning following re-occurring complex auditory patterns in auditory cortical regions.

a) We observed decreased amplitude of activity induced by the repeated acoustic stimuli (reference sequences) compared to those presented only once (fresh sequences). This result was restricted to the beta frequency band in both primary and higher-order regions of the auditory cortex.

Thus, both primary and higher-order auditory cortical regions are actively involved in implicit sensory learning of complex auditory patterns.

References

Agus, T.R., Thorpe, S.J., and Pressnitzer, D. (2010). Rapid Formation of Robust Auditory Memories: Insights from Noise. *Neuron* 66(4), 610-618. doi: 10.1016/j.neuron.2010.04.014.

Ahmed, M., Mallo, T., Leppanen, P.H., Hamalainen, J., Ayravainen, L., Ruusuvirta, T., et al. (2011). Mismatch brain response to speech sound changes in rats. *Front Psychol* 2, 283. doi: 10.3389/fpsyg.2011.00283.

Aitkin, L., Tran, L., and Syka, J. (1994). The responses of neurons in subdivisions of the inferior colliculus of cats to tonal, noise and vocal stimuli. *Exp Brain Res* 98(1), 53-64. doi: 10.1007/BF00229109.

Aitkin, L.M., Anderson, D.J., and Brugge, J.F. (1970). Tonotopic organization and discharge characteristics of single neurons in nuclei of the lateral lemniscus of the cat. *J Neurophysiol* 33(3), 421-440. doi: 10.1152/jn.1970.33.3.421.

Aitkin, L.M., Merzenich, M.M., Irvine, D.R., Clarey, J.C., and Nelson, J.E. (1986). Frequency representation in auditory cortex of the common marmoset (Callithrix jacchus jacchus). *J Comp Neurol* 252(2), 175-185. doi: 10.1002/cne.902520204.

Aitkin, L.M., and Reynolds, A. (1975). Development of binaural responses in the kitten inferior colliculus. *Neurosci Lett* 1(6), 315-319. doi: 10.1016/0304-3940(75)90019-1.

Alho, K., Winkler, I., Escera, C., Huotilainen, M., Virtanen, J., Jaaskelainen, I.P., et al. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. *Psychophysiology* 35(2), 211-224.

Altmann, C.F., Terada, S., Kashino, M., Goto, K., Mima, T., Fukuyama, H., et al. (2014a). Independent or integrated processing of interaural time and level differences in human auditory cortex? *Hear Res* 312, 121-127. doi: 10.1016/j.heares.2014.03.009.

Altmann, C.F., Ueda, R., Furukawa, S., Kashino, M., Mima, T., and Fukuyama, H. (2017). Auditory Mismatch Negativity in Response to Changes of Counter-Balanced Interaural Time and Level Differences. *Front Neurosci* 11, 387. doi: 10.3389/fnins.2017.00387.

Altmann, C.F., Uesaki, M., Ono, K., Matsuhashi, M., Mima, T., and Fukuyama, H. (2014b). Categorical speech perception during active discrimination of consonants and vowels. *Neuropsychologia* 64, 13-23. doi: 10.1016/j.neuropsychologia.2014.09.006.

An, H., Auksztulewicz, R., Kang, H., and Schnupp, J.W.H. (2020). Cortical mapping of mismatch responses to independent acoustic features. *Hear Res*, 107894. doi: 10.1016/j.heares.2020.107894.

Anderson, L.A., Christianson, G.B., and Linden, J.F. (2009). Stimulus-specific adaptation occurs in the auditory thalamus. *J Neurosci* 29(22), 7359-7363. doi: 10.1523/JNEUROSCI.0793-09.2009.

Andrillon, T., Kouider, S., Agus, T., and Pressnitzer, D. (2015). Perceptual learning of acoustic noise generates memory-evoked potentials. *Curr Biol* 25(21), 2823-2829. doi: 10.1016/j.cub.2015.09.027.

Andrillon, T., Pressnitzer, D., Leger, D., and Kouider, S. (2017). Formation and suppression of acoustic memories during human sleep. *Nat Commun* 8(1), 179. doi: 10.1038/s41467-017-00071-z.

Antunes, F.M., Nelken, I., Covey, E., and Malmierca, M.S. (2010). Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *PLoS One* 5(11), e14071. doi: 10.1371/journal.pone.0014071.

Astikainen, P., Ruusuvirta, T., Wikgren, J., and Penttonen, M. (2006). Memory-based detection of rare sound feature combinations in anesthetized rats. *Neuroreport* 17(14), 1561-1564. doi: 10.1097/01.wnr.0000233097.13032.7d.

Astikainen, P., Stefanics, G., Nokia, M., Lipponen, A., Cong, F., Penttonen, M., et al. (2011). Memory-based mismatch response to frequency changes in rats. *PLoS One* 6(9), e24208. doi: 10.1371/journal.pone.0024208.

Auksztulewicz, R., and Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex* 80, 125-140. doi: 10.1016/j.cortex.2015.11.024.

Auksztulewicz, R., Friston, K.J., and Nobre, A.C. (2017). Task relevance modulates the behavioural and neural effects of sensory predictions. *PLoS Biol* 15(12), e2003143. doi: 10.1371/journal.pbio.2003143.

Auksztulewicz, R., Schwiedrzik, C.M., Thesen, T., Doyle, W., Devinsky, O., Nobre, A.C., et al. (2018a). Not All Predictions Are Equal: "What" and "When " Predictions Modulate Activity in Auditory Cortex through Different Mechanisms. *Journal of Neuroscience* 38(40), 8680-8693. doi: 10.1523/Jneurosci.0369-18.2018.

Auksztulewicz, R., Schwiedrzik, C.M., Thesen, T., Doyle, W., Devinsky, O., Nobre, A.C., et al. (2018b). Not All Predictions Are Equal: "What" and "When" Predictions Modulate Activity in Auditory Cortex through Different Mechanisms. *J Neurosci* 38(40), 8680-8693. doi: 10.1523/JNEUROSCI.0369-18.2018.

Bakker, R., Tiesinga, P., and Kotter, R. (2015). The Scalable Brain Atlas: Instant Web-Based Access to Public Brain Atlases and Related Content. *Neuroinformatics* 13(3), 353-366. doi: 10.1007/s12021-014-9258-x.

Baldeweg, T., Klugman, A., Gruzelier, J., and Hirsch, S.R. (2004). Mismatch negativity potentials and cognitive impairment in schizophrenia. *Schizophr Res* 69(2-3), 203-217. doi: 10.1016/j.schres.2003.09.009.

Bao, S., Chang, E.F., Woods, J., and Merzenich, M.M. (2004). Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. *Nat Neurosci* 7(9), 974-981. doi: 10.1038/nn1293.

Bar-Yosef, O., Rotman, Y., and Nelken, I. (2002). Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. *J Neurosci* 22(19), 8619-8632.

Barascud, N., Pearce, M.T., Griffiths, T.D., Friston, K.J., and Chait, M. (2016). Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proc Natl Acad Sci U S A* 113(5), E616-625. doi: 10.1073/pnas.1508523113.

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., and Friston, K.J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76(4), 695-711. doi: 10.1016/j.neuron.2012.10.038.

Bekinschtein, T.A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., and Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proc Natl Acad Sci U S A* 106(5), 1672-1677. doi: 10.1073/pnas.0809667106.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 57(1), 289-300. doi: DOI 10.1111/j.2517-6161.1995.tb02031.x.

Berti, S., Schröger, E., and Mecklinger, A. (2000). Attentive and pre-attentive periodicity analysis in auditory memory: an event-related brain potential study. *Neuroreport* 11(9), 1883-1887. doi: 10.1097/00001756-200006260-00016.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* 10(5), 512-528. doi: 10.1093/cercor/10.5.512.

Bizley, J.K., Walker, K.M.M., Silverman, B.W., King, A.J., and Schnupp, J.W.H. (2009). Interdependent Encoding of Pitch, Timbre, and Spatial Location in Auditory Cortex. *Journal of Neuroscience* 29(7), 2064-2075. doi: 10.1523/Jneurosci.4755-08.2009.

Boly, M., Garrido, M.I., Gosseries, O., Bruno, M.A., Boveroux, P., Schnakers, C., et al. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science* 332(6031), 858-862. doi: 10.1126/science.1202043.

Brawer, J.R., Morest, D.K., and Kane, E.C. (1974). The neuronal architecture of the cochlear nucleus of the cat. *J Comp Neurol* 155(3), 251-300. doi: 10.1002/cne.901550302.

Caclin, A., Brattico, E., Tervaniemi, M., Näätänen, R., Morlet, D., Giard, M.H., et al. (2006). Separate neural processing of timbre dimensions in auditory sensory memory. *J Cogn Neurosci* 18(12), 1959-1972. doi: 10.1162/jocn.2006.18.12.1959.

Camalier, C.R., D'Angelo, W.R., Sterbing-D'Angelo, S.J., de la Mothe, L.A., and Hackett, T.A. (2012). Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *Proceedings of the National Academy of Sciences of the United States of America* 109(44), 18168-18173. doi: 10.1073/pnas.1206387109.

Cant, N.B., and Benson, C.G. (2003). Parallel auditory pathways: projection patterns of the different neuronal populations in the dorsal and ventral cochlear nuclei. *Brain Res Bull* 60(5-6), 457-474. doi: 10.1016/s0361-9230(03)00050-9.

Capsius, B., and Leppelsack, H.J. (1996). Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. *Hear Res* 96(1-2), 59-70. doi: 10.1016/0378-5955(96)00038-x.

Carbajal, G.V., and Malmierca, M.S. (2018). The Neuronal Basis of Predictive Coding Along the Auditory Pathway: From the Subcortical Roots to Cortical Deviance Detection. *Trends Hear* 22, 2331216518784822. doi: 10.1177/2331216518784822.

Casado-Roman, L., Carbajal, G.V., Perez-Gonzalez, D., and Malmierca, M.S. (2020). Prediction error signaling explains neuronal mismatch responses in the medial prefrontal cortex. *PLoS Biol* 18(12), e3001019. doi: 10.1371/journal.pbio.3001019.

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibanez, A., et al. (2013). Expectation and attention in hierarchical auditory prediction. *J Neurosci* 33(27), 11194-11205. doi: 10.1523/JNEUROSCI.0114-13.2013.

Christianson, G.B., Chait, M., de Cheveigne, A., and Linden, J.F. (2014). Auditory evoked fields measured noninvasively with small-animal MEG reveal rapid repetition suppression in the guinea pig. *J Neurophysiol* 112(12), 3053-3065. doi: 10.1152/jn.00189.2014.

Clerici, W.J., and Coleman, J.R. (1990). Anatomy of the rat medial geniculate body: I. Cytoarchitecture, myeloarchitecture, and neocortical connectivity. *J Comp Neurol* 297(1), 14-31. doi: 10.1002/cne.902970103.

Cosmides, L., Barrett, H.C., and Tooby, J. (2010). Colloquium paper: adaptive specializations, social exchange, and the evolution of human intelligence. *Proc Natl Acad Sci U S A* 107 Suppl 2, 9007-9014. doi: 10.1073/pnas.0914623107.

Cosmides, L., and Tooby, J. (2013). Evolutionary psychology: new perspectives on cognition and motivation. *Annu Rev Psychol* 64, 201-229. doi: 10.1146/annurev.psych.121208.131628.

Costa-Faidella, J., Grimm, S., Slabu, L., Diaz-Santaella, F., and Escera, C. (2011). Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. *Psychophysiology* 48(6), 774-783. doi: 10.1111/j.1469-8986.2010.01144.x.

Covey, E., Vater, M., and Casseday, J.H. (1991). Binaural properties of single units in the superior olivary complex of the mustached bat. *J Neurophysiol* 66(3), 1080-1094. doi: 10.1152/jn.1991.66.3.1080.

Curto, C., Sakata, S., Marguet, S., Itskov, V., and Harris, K.D. (2009). A simple model of cortical dynamics explains variability and state dependence of sensory responses in urethane-anesthetized auditory cortex. *J Neurosci* 29(34), 10600-10612. doi: 10.1523/JNEUROSCI.2053-09.2009.

Darwin, C.J., and Hukin, R.W. (2000). Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *J Acoust Soc Am* 107(2), 970-977. doi: 10.1121/1.428278.

David, O., Kilner, J.M., and Friston, K.J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *Neuroimage* 31(4), 1580-1591. doi: 10.1016/j.neuroimage.2006.02.034.

de Cheveigne, A., and Arzounian, D. (2018). Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data. *Neuroimage* 172, 903-912. doi: 10.1016/j.neuroimage.2018.01.035.

De Maesschalck, R., Jouan-Rimbaud, D., and Massart, D.L. (2000). The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems* 50(1), 1-18. doi: Doi 10.1016/S0169-7439(99)00047-7.

Deouell, L.Y., and Bentin, S. (1998). Variable cerebral responses to equally distinct deviance in four auditory dimensions: a mismatch negativity study. *Psychophysiology* 35(6), 745-754.

Doeller, C.F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., and Schröger, E. (2003). Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. *Neuroimage* 20(2), 1270-1282. doi: 10.1016/S1053-8119(03)00389-6.

Doron, N.N., Ledoux, J.E., and Semple, M.N. (2002). Redefining the tonotopic core of rat auditory cortex: physiological evidence for a posterior field. *J Comp Neurol* 453(4), 345-360. doi: 10.1002/cne.10412.

Duque, D., Perez-Gonzalez, D., Ayala, Y.A., Palmer, A.R., and Malmierca, M.S. (2012). Topographic distribution, frequency, and intensity dependence of stimulus-specific adaptation in the inferior colliculus of the rat. *J Neurosci* 32(49), 17762-17774. doi: 10.1523/JNEUROSCI.3190-12.2012.

Duque, D., Wang, X., Nieto-Diego, J., Krumbholz, K., and Malmierca, M.S. (2016). Neurons in the inferior colliculus of the rat show stimulus-specific adaptation for frequency, but not for intensity. *Scientific Reports* 6. doi: ARTN 2411410.1038/srep24114.

Edwards, E., Soltani, M., Deouell, L.Y., Berger, M.S., and Knight, R.T. (2005). High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. *J Neurophysiol* 94(6), 4269-4280. doi: 10.1152/jn.00324.2005.

El Karoui, I., King, J.R., Sitt, J., Meyniel, F., Van Gaal, S., Hasboun, D., et al. (2015). Event-Related Potential, Time-frequency, and Functional Connectivity Facets of Local and

Global Auditory Novelty Processing: An Intracranial Study in Humans. *Cereb Cortex* 25(11), 4203-4212. doi: 10.1093/cercor/bhu143.

Eriksson, J., and Villa, A.E. (2005). Event-related potentials in an auditory oddball situation in the rat. *Biosystems* 79(1-3), 207-212. doi: 10.1016/j.biosystems.2004.09.017.

Escera, C., and Malmierca, M.S. (2014). The auditory novelty system: an attempt to integrate human and animal research. *Psychophysiology* 51(2), 111-123. doi: 10.1111/psyp.12156.

Faber, E.S., and Sah, P. (2003). Calcium-activated potassium channels: multiple contributions to neuronal function. *Neuroscientist* 9(3), 181-194. doi: 10.1177/1073858403009003011.

Farley, B.J., Quirk, M.C., Doherty, J.J., and Christian, E.P. (2010). Stimulus-Specific Adaptation in Auditory Cortex Is an NMDA-Independent Process Distinct from the Sensory Novelty Encoded by the Mismatch Negativity. *Journal of Neuroscience* 30(49), 16475-16484. doi: 10.1523/Jneurosci.2793-10.2010.

Fishman, Y.I., and Steinschneider, M. (2012). Searching for the mismatch negativity in primary auditory cortex of the awake monkey: deviance detection or stimulus specific adaptation? *J Neurosci* 32(45), 15747-15758. doi: 10.1523/JNEUROSCI.2835-12.2012.

Friston, K. (2005). A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360(1456), 815-836. doi: 10.1098/rstb.2005.1622.

Friston, K., and Buzsaki, G. (2016). The Functional Anatomy of Time: What and When in the Brain. *Trends Cogn Sci* 20(7), 500-511. doi: 10.1016/j.tics.2016.05.001.

Gaese, B.H., and Ostwald, J. (2001). Anesthesia changes frequency tuning of neurons in the rat primary auditory cortex. *J Neurophysiol* 86(2), 1062-1066. doi: 10.1152/jn.2001.86.2.1062.

Garrido, M.I., Friston, K.J., Kiebel, S.J., Stephan, K.E., Baldeweg, T., and Kilner, J.M. (2008). The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage* 42(2), 936-944. doi: 10.1016/j.neuroimage.2008.05.018.

Garrido, M.I., Kilner, J.M., Kiebel, S.J., Stephan, K.E., Baldeweg, T., and Friston, K.J. (2009a). Repetition suppression and plasticity in the human brain. *Neuroimage* 48(1), 269-279. doi: 10.1016/j.neuroimage.2009.06.034.

Garrido, M.I., Kilner, J.M., Stephan, K.E., and Friston, K.J. (2009b). The mismatch negativity: a review of underlying mechanisms. *Clin Neurophysiol* 120(3), 453-463. doi: 10.1016/j.clinph.2008.11.029.

Gavornik, J.P., and Bear, M.F. (2014). Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nat Neurosci* 17(5), 732-737. doi: 10.1038/nn.3683.

Giard, M.H., Lavikahen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., et al. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: an event-related potential and dipole-model analysis. *J Cogn Neurosci* 7(2), 133-143. doi: 10.1162/jocn.1995.7.2.133.

Giard, M.H., Perrin, F., Pernier, J., and Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. *Psychophysiology* 27(6), 627-640. doi: 10.1111/j.1469-8986.1990.tb03184.x.

Graybiel, A.M. (1973). The thalamo-cortical projection of the so-called posterior nuclear group: a study with anterograde degeneration methods in the cat. *Brain Res* 49(2), 229-244. doi: 10.1016/0006-8993(73)90420-4.

Grimm, S., Escera, C., Slabu, L., and Costa-Faidella, J. (2011). Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain. *Psychophysiology* 48(3), 377-384. doi: 10.1111/j.1469-8986.2010.01073.x.

Grimm, S., Widmann, A., and Schröger, E. (2004). Differential processing of duration changes within short and long sounds in humans. *Neurosci Lett* 356(2), 83-86. doi: 10.1016/j.neulet.2003.11.035.

Grothe, B., Pecka, M., and McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiol Rev* 90(3), 983-1012. doi: 10.1152/physrev.00026.2009.

Guggenmos, M., Sterzer, P., and Cichy, R.M. (2018). Corrigendum to "Multivariate pattern analysis for MEG: A comparison of dissimilarity measures" [NeuroImage 173 (2018) 434-447]. *Neuroimage* 181, 845. doi: 10.1016/j.neuroimage.2018.07.012.

Guinan, J.J., Jr., and Li, R.Y. (1990). Signal processing in brainstem auditory neurons which receive giant endings (calyces of Held) in the medial nucleus of the trapezoid body of the cat. *Hear Res* 49(1-3), 321-334. doi: 10.1016/0378-5955(90)90111-2.

Haenschel, C., Baldeweg, T., Croft, R.J., Whittington, M., and Gruzelier, J. (2000). Gamma and beta frequency oscillations in response to novel auditory stimuli: A comparison of human electroencephalogram (EEG) data with in vitro models. *Proc Natl Acad Sci U S A* 97(13), 7645-7650. doi: 10.1073/pnas.120162397.

Haenschel, C., Vernon, D.J., Dwivedi, P., Gruzelier, J.H., and Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *J Neurosci* 25(45), 10494-10501. doi: 10.1523/JNEUROSCI.1227-05.2005.

Hara, K., and Harris, R.A. (2002). The anesthetic mechanism of urethane: the effects on neurotransmitter-gated ion channels. *Anesth Analg* 94(2), 313-318, table of contents. doi: 10.1097/00000539-200202000-00015.

Hari, R., Hamalainen, M., Ilmoniemi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., et al. (1984). Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: neuromagnetic recordings in man. *Neurosci Lett* 50(1-3), 127-132. doi: 10.1016/0304-3940(84)90474-9.

Harms, L. (2016). Mismatch responses and deviance detection in N-methyl-D-aspartate (NMDA) receptor hypofunction and developmental models of schizophrenia. *Biol Psychol* 116, 75-81. doi: 10.1016/j.biopsycho.2015.06.015.

Harms, L., Fulham, W.R., Todd, J., Budd, T.W., Hunter, M., Meehan, C., et al. (2014). Mismatch Negativity (MMN) in Freely-Moving Rats with Several Experimental Controls. *Plos One* 9(10). doi: ARTN e11089210.1371/journal.pone.0110892.

Harrington, I.A., Stecker, G.C., Macpherson, E.A., and Middlebrooks, J.C. (2008). Spatial sensitivity of neurons in the anterior, posterior, and primary fields of cat auditory cortex. *Hear Res* 240(1-2), 22-41. doi: 10.1016/j.heares.2008.02.004.

Harris, J.A., Petersen, R.S., and Diamond, M.E. (1999). Distribution of tactile learning and its neural basis. *Proc Natl Acad Sci U S A* 96(13), 7587-7591. doi: 10.1073/pnas.96.13.7587.

Hartmann, T., Schlee, W., and Weisz, N. (2012). It's only in your head: expectancy of aversive auditory stimulation modulates stimulus-induced auditory cortical alpha desynchronization. *Neuroimage* 60(1), 170-178. doi: 10.1016/j.neuroimage.2011.12.034.

Heilbron, M., and Chait, M. (2018). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience* 389, 54-73. doi: 10.1016/j.neuroscience.2017.07.061.

Herbert, H., Aschoff, A., and Ostwald, J. (1991). Topography of projections from the auditory cortex to the inferior colliculus in the rat. *J Comp Neurol* 304(1), 103-122. doi: 10.1002/cne.903040108.

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci* 8(5), 393-402. doi: 10.1038/nrn2113.

Higgins, N.C., Storace, D.A., Escabi, M.A., and Read, H.L. (2010). Specialization of binaural responses in ventral auditory cortices. *J Neurosci* 30(43), 14522-14532. doi: 10.1523/JNEUROSCI.2561-10.2010.

Horvath, J., and Winkler, I. (2004). How the human auditory system treats repetition amongst change. *Neuroscience Letters* 368(2), 157-161. doi: 10.1016/j.neulet.2004.07.004.

Hu, B. (2003). Functional organization of lemniscal and nonlemniscal auditory thalamus. *Exp Brain Res* 153(4), 543-549. doi: 10.1007/s00221-003-1611-5.

Huetz, C., Philibert, B., and Edeline, J.M. (2009). A spike-timing code for discriminating conspecific vocalizations in the thalamocortical system of anesthetized and awake guinea pigs. *J Neurosci* 29(2), 334-350. doi: 10.1523/JNEUROSCI.3269-08.2009.

Hughes, H.C., Darcey, T.M., Barkan, H.I., Williamson, P.D., Roberts, D.W., and Aslin, C.H. (2001). Responses of human auditory association cortex to the omission of an expected acoustic event. *Neuroimage* 13(6 Pt 1), 1073-1089. doi: 10.1006/nimg.2001.0766.

Ille, N., Berg, P., and Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology* 19(2), 113-124. doi: Doi 10.1097/00004691-200203000-00002.

Imada, T., Hari, R., Loveless, N., McEvoy, L., and Sams, M. (1993). Determinants of the auditory mismatch response. *Electroencephalogr Clin Neurophysiol* 87(3), 144-153. doi: 10.1016/0013-4694(93)90120-k.

Imig, T.J., Ruggero, M.A., Kitzes, L.M., Javel, E., and Brugge, J.F. (1977). Organization of auditory cortex in the owl monkey (Aotus trivirgatus). *J Comp Neurol* 171(1), 111-128. doi: 10.1002/cne.901710108.

Ives, D.T., Smith, D.R., and Patterson, R.D. (2005). Discrimination of speaker size from syllable phrases. *J Acoust Soc Am* 118(6), 3816-3822. doi: 10.1121/1.2118427.

Jaaskelainen, I.P., Ahveninen, J., Bonmassar, G., Dale, A.M., Ilmoniemi, R.J., Levanen, S., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proc Natl Acad Sci U S A* 101(17), 6809-6814. doi: 10.1073/pnas.0303760101.

Jaramillo, M., Paavilainen, P., and Näätänen, R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters* 290(2), 101-104. doi: Doi 10.1016/S0304-3940(00)01344-6.

Javitt, D.C., Schroeder, C.E., Steinschneider, M., Arezzo, J.C., and Vaughan, H.G., Jr. (1992). Demonstration of mismatch negativity in the monkey. *Electroencephalogr Clin Neurophysiol* 83(1), 87-90. doi: 10.1016/0013-4694(92)90137-7.

Javitt, D.C., Steinschneider, M., Schroeder, C.E., and Arezzo, J.C. (1996). Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. *Proc Natl Acad Sci U S A* 93(21), 11962-11967. doi: 10.1073/pnas.93.21.11962.

Jenkins, W.M., and Masterton, R.B. (1982). Sound localization: effects of unilateral lesions in central auditory system. *J Neurophysiol* 47(6), 987-1016. doi: 10.1152/jn.1982.47.6.987.

Jones, E.G. (2003). Chemically defined parallel pathways in the monkey auditory system. *Ann N Y Acad Sci* 999, 218-233. doi: 10.1196/annals.1284.033.

Joutsiniemi, S.L., Ilvonen, T., Sinkkonen, J., Huotilainen, M., Tervaniemi, M., Lehtokoski, A., et al. (1998). The mismatch negativity for duration decrement of auditory stimuli in healthy

subjects. *Electroencephalogr Clin Neurophysiol* 108(2), 154-159. doi: 10.1016/s0168-5597(97)00082-8.
Kaas, J.H., and Hackett, T.A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A* 97(22), 11793-11799. doi: 10.1073/pnas.97.22.11793.
Kang, H., Agus, T.R., and Pressnitzer, D. (2017). Auditory memory for random time patterns. *J Acoust Soc Am* 142(4), 2219. doi: 10.1121/1.5007730.
Kavanagh, G.L., and Kelly, J.B. (1992). Midline and Lateral Field Sound Localization in the Ferret (Mustela-Putorius) - Contribution of the Superior Olivary Complex. *Journal of Neurophysiology* 67(6), 1643-1658.
Kawahara, H. (2006). STRAIGHT, exploitation of the other aspect of VOCODER: Perceptually isomorphic decomposition of speech sounds. *Acoustical Science and Technology* 27(6), 349-353. doi: 10.1250/ast.27.349.
Kiebel, S.J., Daunizeau, J., and Friston, K.J. (2008). A hierarchy of time-scales and the brain. *PLoS Comput Biol* 4(11), e1000209. doi: 10.1371/journal.pcbi.1000209.
Klimesch, W., Russegger, H., Doppelmayr, M., and Pachinger, T. (1998). A method for the calculation of induced band power: implications for the significance of brain oscillations. *Electroencephalogr Clin Neurophysiol* 108(2), 123-130. doi: 10.1016/s0168-5597(97)00078-6.
Komatsu, M., Takaura, K., and Fujii, N. (2015). Mismatch negativity in common marmosets: Whole-cortical recordings with multi-channel electrocorticograms. *Sci Rep* 5, 15006. doi: 10.1038/srep15006.
Korzyukov, O., Alho, K., Kujala, A., Gumenyuk, V., Ilmoniemi, R.J., Virtanen, J., et al. (1999). Electromagnetic responses of the human auditory cortex generated by sensory-memory based processing of tone-frequency changes. *Neurosci Lett* 276(3), 169-172. doi: 10.1016/s0304-3940(99)00807-1.
Korzyukov, O.A., Winkler, I., Gumenyuk, V.I., and Alho, K. (2003). Processing abstract auditory features in the human auditory cortex. *Neuroimage* 20(4), 2245-2258. doi: 10.1016/j.neuroimage.2003.08.014.
Kosaki, H., Hashikawa, T., He, J., and Jones, E.G. (1997). Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. *J Comp Neurol* 386(2), 304-316.
Kosmal, A., Malinowska, M., and Kowalska, D.M. (1997). Thalamic and amygdaloid connections of the auditory association cortex of the superior temporal gyrus in rhesus monkey (Macaca mulatta). *Acta Neurobiol Exp (Wars)* 57(3), 165-188.
Kraus, N., McGee, T., Carrell, T., King, C., Littman, T., and Nicol, T. (1994). Discrimination of speech-like contrasts in the auditory thalamus and cortex. *J Acoust Soc Am* 96(5 Pt 1), 2758-2768. doi: 10.1121/1.411282.
Kujala, T., and Näätänen, R. (2001). The mismatch negativity in evaluating central auditory dysfunction in dyslexia. *Neurosci Biobehav Rev* 25(6), 535-543. doi: 10.1016/s0149-7634(01)00032-x.
Kusmierek, P., and Rauschecker, J.P. (2009). Functional specialization of medial auditory belt cortex in the alert rhesus monkey. *J Neurophysiol* 102(3), 1606-1622. doi: 10.1152/jn.00167.2009.
Kusmierek, P., and Rauschecker, J.P. (2014). Selectivity for space and time in early areas of the auditory dorsal stream in the rhesus monkey. *J Neurophysiol* 111(8), 1671-1685. doi: 10.1152/jn.00436.2013.

Lamas, V., Estevez, S., Pernia, M., Plaza, I., and Merchan, M.A. (2017). Stereotactically-guided Ablation of the Rat Auditory Cortex, and Localization of the Lesion in the Brain. *J Vis Exp* (128). doi: 10.3791/56429.

Lan, T., Erdogmus, D., Black, L., and Van Santen, J. (2010). A comparison of different dimensionality reduction and feature selection methods for single trial ERP detection. *Annu Int Conf IEEE Eng Med Biol Soc* 2010, 6329-6332. doi: 10.1109/IEMBS.2010.5627642.

Lazar, R., and Metherate, R. (2003). Spectral interactions, but no mismatch negativity, in auditory cortex of anesthetized rat. *Hearing Research* 181(1-2), 51-56. doi: 10.1016/S0378-5955(03)00166-7.

Lee, C.C., and Winer, J.A. (2008). Connections of cat auditory cortex: I. Thalamocortical system. *J Comp Neurol* 507(6), 1879-1900. doi: 10.1002/cne.21611.

Lee, Y.L., Lopez, D.E., Meloni, E.G., and Davis, M. (1996). A primary acoustic startle pathway: Obligatory role of cochlear root neurons and the nucleus reticularis pontis caudalis. *Journal of Neuroscience* 16(11), 3775-3789.

Leung, S., Cornella, M., Grimm, S., and Escera, C. (2012). Is fast auditory change detection feature specific? An electrophysiological study in humans. *Psychophysiology* 49(7), 933-942. doi: 10.1111/j.1469-8986.2012.01375.x.

Li, D., Christ, S.E., and Cowan, N. (2014). Domain-general and domain-specific functional networks in working memory. *Neuroimage* 102 Pt 2, 646-656. doi: 10.1016/j.neuroimage.2014.08.028.

Li, K., Chan, C.H.K., Rajendran, V.G., Meng, Q., Rosskothen-Kuhl, N., and Schnupp, J.W.H. (2019). Microsecond sensitivity to envelope interaural time differences in rats. *J Acoust Soc Am* 145(5), EL341. doi: 10.1121/1.5099164.

Lim, Y., Lagoy, R., Shinn-Cunningham, B.G., and Gardner, T.J. (2016). Transformation of temporal sequences in the zebra finch auditory system. *Elife* 5. doi: 10.7554/eLife.18205.

Lipponen, A., Kurkela, J.L.O., Kylaheiko, I., Holtta, S., Ruusuvirta, T., Hamalainen, J.A., et al. (2019). Auditory-evoked potentials to changes in sound duration in urethane-anaesthetized mice. *Eur J Neurosci* 50(2), 1911-1919. doi: 10.1111/ejn.14359.

Lu, K., Liu, W., Zan, P., David, S.V., Fritz, J.B., and Shamma, S.A. (2018). Implicit Memory for Complex Sounds in Higher Auditory Cortex of the Ferret. *J Neurosci* 38(46), 9955-9966. doi: 10.1523/JNEUROSCI.2118-18.2018.

Lumani, A., and Zhang, H. (2010). Responses of neurons in the rat's dorsal cortex of the inferior colliculus to monaural tone bursts. *Brain Res* 1351, 115-129. doi: 10.1016/j.brainres.2010.06.066.

Luo, H., Tian, X., Song, K., Zhou, K., and Poeppel, D. (2013). Neural Response Phase Tracks How Listeners Learn New Acoustic Representations. *Current Biology* 23(11), 968-974. doi: 10.1016/j.cub.2013.04.031.

Macdonald, M., and Campbell, K. (2011). Effects of a violation of an expected increase or decrease in intensity on detection of change within an auditory pattern. *Brain and Cognition* 77(3), 438-445. doi: 10.1016/j.bandc.2011.08.014.

Mahmoudzadeh, M., Dehaene-Lambertz, G., and Wallois, F. (2017). Electrophysiological and hemodynamic mismatch responses in rats listening to human speech syllables. *PLoS One* 12(3), e0173801. doi: 10.1371/journal.pone.0173801.

Malmierca, M.S. (2003). The structure and physiology of the rat auditory system: an overview. *Int Rev Neurobiol* 56, 147-211. doi: 10.1016/s0074-7742(03)56005-6.

Malmierca, M.S., Anderson, L.A., and Antunes, F.M. (2015). The cortical modulation of stimulus-specific adaptation in the auditory midbrain and thalamus: a potential neuronal correlate for predictive coding. *Front Syst Neurosci* 9, 19. doi: 10.3389/fnsys.2015.00019.

Malmierca, M.S., Blackstad, T.W., and Osen, K.K. (2011). Computer-assisted 3-D reconstructions of Golgi-impregnated neurons in the cortical regions of the inferior colliculus of rat. *Hear Res* 274(1-2), 13-26. doi: 10.1016/j.heares.2010.06.011.

Malmierca, M.S., Blackstad, T.W., Osen, K.K., Karagulle, T., and Molowny, R.L. (1993). The central nucleus of the inferior colliculus in rat: a Golgi and computer reconstruction study of neuronal and laminar structure. *J Comp Neurol* 333(1), 1-27. doi: 10.1002/cne.903330102.

Malmierca, M.S., Cristaudo, S., Perez-Gonzalez, D., and Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J Neurosci* 29(17), 5483-5493. doi: 10.1523/JNEUROSCI.4153-08.2009.

Malmierca, M.S., and Ryugo, D.K. (2011). Descending Connections of Auditory Cortex to the Midbrain and Brain Stem. *Auditory Cortex*, 189-208. doi: 10.1007/978-1-4419-0074-6_9.

Malmierca, M.S., Sanchez-Vives, M.V., Escera, C., and Bendixen, A. (2014). Neuronal adaptation, novelty detection and regularity encoding in audition. *Front Syst Neurosci* 8, 111. doi: 10.3389/fnsys.2014.00111.

Malmierca, M.S., Seip, K.L., and Osen, K.K. (1995). Morphological classification and identification of neurons in the inferior colliculus: a multivariate analysis. *Anat Embryol (Berl)* 191(4), 343-350. doi: 10.1007/BF00534687.

May, P., Tiitinen, H., Ilmoniemi, R.J., Nyman, G., Taylor, J.G., and Näätänen, R. (1999). Frequency change detection in human auditory cortex. *J Comput Neurosci* 6(2), 99-120. doi: 10.1023/a:1008896417606.

May, P.J., and Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology* 47(1), 66-122. doi: 10.1111/j.1469-8986.2009.00856.x.

Merzenich, M.M., and Brugge, J.F. (1973). Representation of the cochlear partition of the superior temporal plane of the macaque monkey. *Brain Res* 50(2), 275-296. doi: 10.1016/0006-8993(73)90731-2.

Moran, R.J., Campo, P., Symmonds, M., Stephan, K.E., Dolan, R.J., and Friston, K.J. (2013). Free energy, precision and learning: the role of cholinergic neuromodulation. *J Neurosci* 33(19), 8227-8236. doi: 10.1523/JNEUROSCI.4255-12.2013.

Morel, A., Garraghty, P.E., and Kaas, J.H. (1993). Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *J Comp Neurol* 335(3), 437-459. doi: 10.1002/cne.903350312.

Morel, A., and Kaas, J.H. (1992). Subdivisions and connections of auditory cortex in owl monkeys. *J Comp Neurol* 318(1), 27-63. doi: 10.1002/cne.903180104.

Morest, D.K., and Oliver, D.L. (1984). The neuronal architecture of the inferior colliculus in the cat: defining the functional anatomy of the auditory midbrain. *J Comp Neurol* 222(2), 209-236. doi: 10.1002/cne.902220206.

Movshon, J.A., and Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. *Nature* 278(5707), 850-852. doi: 10.1038/278850a0.

Näätänen, R. (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), 201-232. doi: Doi 10.1017/S0140525x00078407.

Näätänen, R. (1995). The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear Hear* 16(1), 6-18.

Näätänen, R. (2007). The mismatch negativity - Where is the big fish? *Journal of Psychophysiology* 21(3-4), 133-137. doi: 10.1027/0269-8803.21.34.133.

Näätänen, R. (1992). *Attention and brain function.* Hillsdale, N.J.: L. Erlbaum.

Näätänen, R., and Alho, K. (1995). Generators of electrical and magnetic mismatch responses in humans. *Brain Topogr* 7(4), 315-320. doi: 10.1007/BF01195257.

Näätänen, R., Jacobsen, T., and Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. *Psychophysiology* 42(1), 25-32. doi: 10.1111/j.1469-8986.2005.00256.x.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385(6615), 432-434. doi: 10.1038/385432a0.

Näätänen, R., and Michie, P.T. (1979). Early selective-attention effects on the evoked potential: a critical review and reinterpretation. *Biol Psychol* 8(2), 81-136. doi: 10.1016/0301-0511(79)90053-x.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 118(12), 2544-2590. doi: 10.1016/j.clinph.2007.04.026.

Näätänen, R., and Rinne, T. (2002). Electric brain response to sound repetition in humans: an index of long-term-memory - trace formation? *Neurosci Lett* 318(1), 49-51. doi: 10.1016/s0304-3940(01)02438-7.

Nakamura, T., Michie, P.T., Fulham, W.R., Todd, J., Budd, T.W., Schell, U., et al. (2011). Epidural auditory event-related potentials in the rat to frequency and duration deviants: evidence of mismatch negativity? *Frontiers in Psychology* 2. doi: ARTN 367 10.3389/fpsyg.2011.00367.

Netser, S., Zahar, Y., and Gutfreund, Y. (2011). Stimulus-Specific Adaptation: Can It Be a Neural Correlate of Behavioral Habituation? *Journal of Neuroscience* 31(49), 17811-17820. doi: 10.1523/Jneurosci.4790-11.2011.

Nieto-Diego, J., and Malmierca, M.S. (2016). Topographic Distribution of Stimulus-Specific Adaptation across Auditory Cortical Fields in the Anesthetized Rat. *PLoS Biol* 14(3), e1002397. doi: 10.1371/journal.pbio.1002397.

Noda, T., and Takahashi, H. (2015). Anesthetic effects of isoflurane on the tonotopic map and neuronal population activity in the rat auditory cortex. *Eur J Neurosci* 42(6), 2298-2311. doi: 10.1111/ejn.13007.

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D.Y., and Schröger, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage* 15(1), 167-174. doi: 10.1006/nimg.2001.0970.

Osen, K.K. (1969). Cytoarchitecture of the cochlear nuclei in the cat. *J Comp Neurol* 136(4), 453-484. doi: 10.1002/cne.901360407.

Paavilainen, P., Valppu, S., and Näätänen, R. (2001). The additivity of the auditory feature analysis in the human brain as indexed by the mismatch negativity: 1+1 approximate to 2 but 1+1+1 < 3. *Neuroscience Letters* 301(3), 179-182. doi: Doi 10.1016/S0304-3940(01)01635-4.

Pagliardini, S., Greer, J.J., Funk, G.D., and Dickson, C.T. (2012). State-dependent modulation of breathing in urethane-anesthetized rats. *J Neurosci* 32(33), 11259-11270. doi: 10.1523/JNEUROSCI.0948-12.2012.

Parras, G.G., Nieto-Diego, J., Carbajal, G.V., Valdes-Baizabal, C., Escera, C., and Malmierca, M.S. (2017). Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nat Commun* 8(1), 2148. doi: 10.1038/s41467-017-02038-6.

Pearce, M.T., Ruiz, M.H., Kapasi, S., Wiggins, G.A., and Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *Neuroimage* 50(1), 302-313. doi: 10.1016/j.neuroimage.2009.12.019.

Perez-Gonzalez, D., Malmierca, M.S., and Covey, E. (2005). Novelty detector neurons in the mammalian auditory midbrain. *European Journal of Neuroscience* 22(11), 2879-2885. doi: 10.1111/j.1460-9568.2005.04472.x.

Phillips, H.N., Blenkmann, A., Hughes, L.E., Bekinschtein, T.A., and Rowe, J.B. (2015). Hierarchical Organization of Frontotemporal Networks for the Prediction of Stimuli across Multiple Dimensions. *J Neurosci* 35(25), 9255-9264. doi: 10.1523/JNEUROSCI.5095-14.2015.

Pincze, Z., Lakatos, P., Rajkai, C., Ulbert, I., and Karmos, G. (2001). Separation of mismatch negativity and the N1 wave in the auditory cortex of the cat: a topographic study. *Clin Neurophysiol* 112(5), 778-784. doi: 10.1016/s1388-2457(01)00509-0.

Polley, D.B., Read, H.L., Storace, D.A., and Merzenich, M.M. (2007). Multiparametric auditory receptive field organization across five cortical fields in the albino rat. *J Neurophysiol* 97(5), 3621-3638. doi: 10.1152/jn.01298.2006.

Rauschecker, J.P., and Scott, S.K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* 12(6), 718-724. doi: 10.1038/nn.2331.

Rauschecker, J.P., and Tian, B. (2004). Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol* 91(6), 2578-2589. doi: 10.1152/jn.00834.2003.

Rauschecker, J.P., Tian, B., and Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268(5207), 111-114. doi: 10.1126/science.7701330.

Rauschecker, J.P., Tian, B., Pons, T., and Mishkin, M. (1997). Serial and parallel processing in rhesus monkey auditory cortex. *J Comp Neurol* 382(1), 89-103.

Recanzone, G.H. (2000). Spatial processing in the auditory cortex of the macaque monkey. *Proc Natl Acad Sci U S A* 97(22), 11829-11835. doi: 10.1073/pnas.97.22.11829.

Recanzone, G.H., Guard, D.C., and Phan, M.L. (2000). Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J Neurophysiol* 83(4), 2315-2331. doi: 10.1152/jn.2000.83.4.2315.

Recasens, M., Grimm, S., Wollbrink, A., Pantev, C., and Escera, C. (2014). Encoding of Nested Levels of Acoustic Regularity in Hierarchically Organized Areas of the Human Auditory Cortex. *Human Brain Mapping* 35(11), 5701-5716. doi: 10.1002/hbm.22582.

Reches, A., and Gutfreund, Y. (2008). Stimulus-specific adaptations in the gaze control system of the barn owl. *J Neurosci* 28(6), 1523-1533. doi: 10.1523/JNEUROSCI.3785-07.2008.

Retsa, C., Matusz, P.J., Schnupp, J.W.H., and Murray, M.M. (2018). What's what in auditory cortices? *Neuroimage* 176, 29-40. doi: 10.1016/j.neuroimage.2018.04.028.

Richardson, B.D., Hancock, K.E., and Caspary, D.M. (2013). Stimulus-specific adaptation in auditory thalamus of young and aged awake rats. *J Neurophysiol* 110(8), 1892-1902. doi: 10.1152/jn.00403.2013.

Rockel, A.J., and Jones, E.G. (1973). The neuronal organization of the inferior colliculus of the adult cat. I. The central nucleus. *J Comp Neurol* 147(1), 11-60. doi: 10.1002/cne.901470103.

Roger, C., Hasbroucq, T., Rabat, A., Vidal, F., and Burle, B. (2009a). Neurophysics of temporal discrimination in the rat: A mismatch negativity study. *Psychophysiology* 46(5), 1028-1032. doi: 10.1111/j.1469-8986.2009.00840.x.

Roger, C., Hasbroucq, T., Rabat, A., Vidal, F., and Burle, B. (2009b). Neurophysics of temporal discrimination in the rat: a mismatch negativity study. *Psychophysiology* 46(5), 1028-1032. doi: 10.1111/j.1469-8986.2009.00840.x.

Romanski, L.M., Giguere, M., Bates, J.F., and Goldman-Rakic, P.S. (1997). Topographic organization of medial pulvinar connections with the prefrontal cortex in the rhesus monkey. *J Comp Neurol* 379(3), 313-332.

Rosburg, T., Weigl, M., Thiel, R., and Mager, R. (2018). The event-related potential component P3a is diminished by identical deviance repetition, but not by non-identical repetitions. *Experimental Brain Research* 236(5), 1519-1530. doi: 10.1007/s00221-018-5237-z.

Rosch, R.E., Auksztulewicz, R., Leung, P.D., Friston, K.J., and Baldeweg, T. (2019). Selective Prefrontal Disinhibition in a Roving Auditory Oddball Paradigm Under N-Methyl-D-Aspartate Receptor Blockade. *Biol Psychiatry Cogn Neurosci Neuroimaging* 4(2), 140-150. doi: 10.1016/j.bpsc.2018.07.003.

Rosenthal, C.R., Andrews, S.K., Antoniades, C.A., Kennard, C., and Soto, D. (2016). Learning and Recognition of a Non-conscious Sequence of Events in Human Primary Visual Cortex. *Curr Biol* 26(6), 834-841. doi: 10.1016/j.cub.2016.01.040.

Ruhnau, P., Herrmann, B., and Schröger, E. (2012). Finding the right control: the mismatch negativity under investigation. *Clin Neurophysiol* 123(3), 507-512. doi: 10.1016/j.clinph.2011.07.035.

Ruusuvirta, T., Penttonen, M., and Korhonen, T. (1998). Auditory cortical event-related potentials to pitch deviances in rats. *Neurosci Lett* 248(1), 45-48. doi: 10.1016/s0304-3940(98)00330-9.

Saarinen, J., Paavilainen, P., Schoger, E., Tervaniemi, M., and Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *Neuroreport* 3(12), 1149-1151. doi: 10.1097/00001756-199212000-00030.

Saldana, E., Feliciano, M., and Mugnaini, E. (1996). Distribution of descending projections from primary auditory neocortex to inferior colliculus mimics the topography of intracollicular projections. *Journal of Comparative Neurology* 371(1), 15-40. doi: Doi 10.1002/(Sici)1096-9861(19960715)371:1<15::Aid-Cne2>3.0.Co;2-O.

Salisbury, D.F. (2012). Finding the missing stimulus mismatch negativity (MMN): emitted MMN to violations of an auditory gestalt. *Psychophysiology* 49(4), 544-548. doi: 10.1111/j.1469-8986.2011.01336.x.

Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr Clin Neurophysiol* 62(6), 437-448. doi: 10.1016/0168-5597(85)90054-1.

Scholz, S., Schneider, S.L., and Rose, M. (2017). Differential effects of ongoing EEG beta and theta power on memory formation. *PLoS One* 12(2), e0171913. doi: 10.1371/journal.pone.0171913.

Schröger, E. (1995). Processing of auditory deviants with changes in one versus two stimulus dimensions. *Psychophysiology* 32(1), 55-65. doi: 10.1111/j.1469-8986.1995.tb03406.x.

Schröger, E. (1996). Interaural time and level differences: integrated or separated processing? *Hear Res* 96(1-2), 191-198. doi: 10.1016/0378-5955(96)00066-4.

Schröger, E., and Wolff, C. (1996). Mismatch response of the human brain to changes in sound location. *Neuroreport* 7(18), 3005-3008. doi: 10.1097/00001756-199611250-00041.

Schubotz, R.I., von Cramon, D.Y., and Lohmann, G. (2003). Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. *Neuroimage* 20(1), 173-185. doi: 10.1016/S1053-8119(03)00218-0.

Schultz, S.K. (2001). Principles of neural science, 4th edition. *American Journal of Psychiatry* 158(4), 662-662. doi: DOI 10.1176/appi.ajp.158.4.662.

Scott, S.K., Blank, C.C., Rosen, S., and Wise, R.J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123 Pt 12, 2400-2406. doi: 10.1093/brain/123.12.2400.

Sharma, A., and Dorman, M.F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *J Acoust Soc Am* 107(5 Pt 1), 2697-2703. doi: 10.1121/1.428655.

Shelley, A.M., Ward, P.B., Catts, S.V., Michie, P.T., Andrews, S., and McConaghy, N. (1991). Mismatch negativity: an index of a preattentive processing deficit in schizophrenia. *Biol Psychiatry* 30(10), 1059-1062. doi: 10.1016/0006-3223(91)90126-7.

Shiramatsu, T.I., Kanzaki, R., and Takahashi, H. (2013). Cortical mapping of mismatch negativity with deviance detection property in rat. *PLoS One* 8(12), e82663. doi: 10.1371/journal.pone.0082663.

Shiramatsu, T.I., and Takahashi, H. (2018). Mismatch Negativity in Rat Auditory Cortex Represents the Empirical Salience of Sounds. *Front Neurosci* 12, 924. doi: 10.3389/fnins.2018.00924.

Skipper, J.I. (2014). Echoes of the spoken past: how auditory cortex hears context during speech perception. *Philos Trans R Soc Lond B Biol Sci* 369(1651), 20130297. doi: 10.1098/rstb.2013.0297.

Sloutsky, V.M. (2010). Mechanisms of cognitive development: domain-general learning or domain-specific constraints? *Cogn Sci* 34(7), 1125-1130. doi: 10.1111/j.1551-6709.2010.01132.x.

Southwell, R., and Chait, M. (2018). Enhanced deviant responses in patterned relative to random sound sequences. *Cortex* 109, 92-103. doi: 10.1016/j.cortex.2018.08.032.

Stecker, G.C., Mickey, B.J., Macpherson, E.A., and Middlebrooks, J.C. (2003). Spatial sensitivity in field PAF of cat auditory cortex. *J Neurophysiol* 89(6), 2889-2903. doi: 10.1152/jn.00980.2002.

Stefanics, G., Stephan, K.E., and Heinzle, J. (2019). Feature-specific prediction errors for visual mismatch. *Neuroimage* 196, 142-151. doi: 10.1016/j.neuroimage.2019.04.020.

Sussman, E., Ritter, W., and Vaughan, H.G., Jr. (1998). Predictability of stimulus deviance and the mismatch negativity. *Neuroreport* 9(18), 4167-4170. doi: 10.1097/00001756-199812210-00031.

Sussman, E., and Winkler, I. (2001). Dynamic sensory updating in the auditory system. *Brain Res Cogn Brain Res* 12(3), 431-439. doi: 10.1016/s0926-6410(01)00067-2.

Sussman, E., Winkler, I., and Schröger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychon Bull Rev* 10(3), 630-637. doi: 10.3758/bf03196525.

Sussman, E.S., and Gumenyuk, V. (2005). Organization of sequential sounds in auditory memory. *Neuroreport* 16(13), 1519-1523. doi: 10.1097/01.wnr.0000177002.35193.4c.

Taaseh, N., Yaron, A., and Nelken, I. (2011). Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS One* 6(8), e23369. doi: 10.1371/journal.pone.0023369.

Takaura, K., and Fujii, N. (2016). Facilitative effect of repetitive presentation of one stimulus on cortical responses to other stimuli in macaque monkeys--a possible neural mechanism for mismatch negativity. *Eur J Neurosci* 43(4), 516-528. doi: 10.1111/ejn.13136.

Takegata, R., Paavilainen, P., Näätänen, R., and Winkler, I. (1999). Independent processing of changes in auditory single features and feature conjunctions in humans as indexed by the mismatch negativity. *Neurosci Lett* 266(2), 109-112. doi: 10.1016/s0304-3940(99)00267-0.

Takegata, R., Tervaniemi, M., Alku, P., Ylinen, S., and Näätänen, R. (2008). Parameter-specific modulation of the mismatch negativity to duration decrement and increment: evidence for asymmetric processes. *Clin Neurophysiol* 119(7), 1515-1523. doi: 10.1016/j.clinph.2008.03.025.

Tardif, E., Murray, M.M., Meylan, R., Spierer, L., and Clarke, S. (2006). The spatio-temporal brain dynamics of processing and integrating sound localization cues in humans. *Brain Res* 1092(1), 161-176. doi: 10.1016/j.brainres.2006.03.095.

Tervaniemi, M., Schröger, E., Saher, M., and Näätänen, R. (2000). Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans - a mismatch negativity study. *Neuroscience Letters* 290(1), 66-70. doi: Doi 10.1016/S0304-3940(00)01290-8.

Tian, B., Reser, D., Durham, A., Kustov, A., and Rauschecker, J.P. (2001). Functional specialization in rhesus monkey auditory cortex. *Science* 292(5515), 290-293. doi: 10.1126/science.1058911.

Tikhonravov, D., Neuvonen, T., Pertovaara, A., Savioja, K., Ruusuvirta, T., Näätänen, R., et al. (2008). Effects of an NMDA-receptor antagonist MK-801 on an MMN-like response recorded in anesthetized rats. *Brain Res* 1203, 97-102. doi: 10.1016/j.brainres.2008.02.006.

Tikhonravov, D., Neuvonen, T., Pertovaara, A., Savioja, K., Ruusuvirta, T., Näätänen, R., et al. (2010). Dose-related effects of memantine on a mismatch negativity-like response in anesthetized rats. *Neuroscience* 167(4), 1175-1182. doi: 10.1016/j.neuroscience.2010.03.014.

Uhrig, L., Dehaene, S., and Jarraya, B. (2014). A hierarchy of responses to auditory regularities in the macaque brain. *J Neurosci* 34(4), 1127-1132. doi: 10.1523/JNEUROSCI.3165-13.2014.

Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004a). Multiple time scales of adaptation in auditory cortex neurons. *Journal of Neuroscience* 24(46), 10440-10453. doi: 10.1523/Jneurosci.1905-04.2004.

Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004b). Multiple time scales of adaptation in auditory cortex neurons. *J Neurosci* 24(46), 10440-10453. doi: 10.1523/JNEUROSCI.1905-04.2004.

Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6(4), 391-398. doi: 10.1038/nn1032.

Umbricht, D., Vyssotki, D., Latanov, A., Nitsch, R., and Lipp, H.P. (2005). Deviance-related electrophysiological activity in mice: is there mismatch negativity in mice? *Clinical Neurophysiology* 116(2), 353-363. doi: 10.1016/j.clinph.2004.08.015.

von der Behrens, W., Bauerle, P., Kossl, M., and Gaese, B.H. (2009). Correlating stimulus-specific adaptation of cortical neurons and local field potentials in the awake rat. *J Neurosci* 29(44), 13837-13849. doi: 10.1523/JNEUROSCI.3475-09.2009.

Wacongne, C., Changeux, J.P., and Dehaene, S. (2012). A Neuronal Model of Predictive Coding Accounting for the Mismatch Negativity. *Journal of Neuroscience* 32(11), 3665-3678. doi: 10.1523/Jneurosci.5003-11.2012.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., and Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proc Natl Acad Sci U S A* 108(51), 20754-20759. doi: 10.1073/pnas.1117807108.

Walker, K.M., Schnupp, J.W., Hart-Schnupp, S.M., King, A.J., and Bizley, J.K. (2009). Pitch discrimination by ferrets for simple and complex sounds. *J Acoust Soc Am* 126(3), 1321-1335. doi: 10.1121/1.3179676.

Wessinger, C.M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., and Rauschecker, J.P. (2001). Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *J Cogn Neurosci* 13(1), 1-7. doi: 10.1162/089892901564108.

Winer, J.A., Kelly, J.B., and Larue, D.T. (1999). Neural architecture of the rat medial geniculate body. *Hear Res* 130(1-2), 19-41. doi: 10.1016/s0378-5955(98)00216-0.

Winer, J.A., Larue, D.T., and Pollak, G.D. (1995). GABA and glycine in the central auditory system of the mustache bat: structural substrates for inhibitory neuronal organization. *J Comp Neurol* 355(3), 317-353. doi: 10.1002/cne.903550302.

Winkler, I., and Czigler, I. (1998). Mismatch negativity: deviance detection or the maintenance of the 'standard'. *Neuroreport* 9(17), 3809-3813. doi: 10.1097/00001756-199812010-00008.

Wolff, M.J., Kandemir, G., Stokes, M.G., and Akyurek, E.G. (2020). Unimodal and Bimodal Access to Sensory Working Memories by Auditory and Visual Impulses. *J Neurosci* 40(3), 671-681. doi: 10.1523/JNEUROSCI.1194-19.2019.

Woods, V., Trumpis, M., Bent, B., Palopoli-Trojani, K., Chiang, C.H., Wang, C., et al. (2018). Long-term recording reliability of liquid crystal polymer mu ECoG arrays. *Journal of Neural Engineering* 15(6). doi: ARTN 06602410.1088/1741-2552/aae39d.

Yeterian, E.H., and Pandya, D.N. (1989). Thalamic connections of the cortex of the superior temporal sulcus in the rhesus monkey. *J Comp Neurol* 282(1), 80-97. doi: 10.1002/cne.902820107.

Yin, T.C., and Chan, J.C. (1990). Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol* 64(2), 465-488. doi: 10.1152/jn.1990.64.2.465.

Yu, X.J., Xu, X.X., He, S., and He, J. (2009). Change detection by thalamic reticular neurons. *Nat Neurosci* 12(9), 1165-1170. doi: 10.1038/nn.2373.

Zhao, L., Liu, Y., Shen, L., Feng, L., and Hong, B. (2011). Stimulus-specific adaptation and its dynamics in the inferior colliculus of rat. *Neuroscience* 181, 163-174. doi: 10.1016/j.neuroscience.2011.01.060.

Zurita, P., Villa, A.E., de Ribaupierre, Y., de Ribaupierre, F., and Rouiller, E.M. (1994). Changes of single unit activity in the cat's auditory thalamus and cortex associated to different anesthetic conditions. *Neurosci Res* 19(3), 303-316. doi: 10.1016/0168-0102(94)90043-4.

Appendix A

Publications

1 Journals

- **HyunJung An,** Ryszard Auksztulewicz, HiJee Kang, Jan W.H. (2020) **Cortical mapping of mismatch responses to independent acoustic features** *Hearing Research* **(**DOI: https://doi.org/10.1016/j.heares.2020.107894)

- **HyunJung An**, Shing Ho Kei, Ryszard Auksztulewicz*, Jan W.H. Schnupp* (2021) Do auditory mismatch responses differ between acoustic features? *Frontiers in Human Neuroscience* (DOI: https://doi.org/10.3389/fnhum.2021.613903)

- HiJee Kang, Ryszard Auksztulewicz, **HyunJung An**, Nicolas Abichacra, Mitchell Sutter, Jan W.H. Schnupp (2021) Neural correlates of auditory pattern learning in the auditory cortex *Frontiers in Neuroscience* (DOI: https://doi: 10.3389/fnins.2021.610978)

2 Poster

- **HyunJung An**, Ryszard Auksztulewicz, HiJee Kang, Jan W.H**. Cortical mapping of prediction error responses to multiple sensory features** 50th Annual Society for Neuroscience's Annual Meeting (SFN 2019)