Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Présentée et soutenue par :

Jayalakshmi VISWANATHAN

le mardi 20 décembre 2016

Titre :

When noise enlightens: an investigation into the mechanisms of perception and long-term memory for meaningless auditory stimuli.

Quand le bruit nous éclaire: une étude sur les mécanismes de la perception et de la mémoire à long-terme pour des stimuli auditifs sans signification.

École doctorale et discipline ou spécialité**:**

ED CLESCO: Neurosciences

Unité de recherche :

CerCo - Centre de recherche cerveau et cognition

Directeur/trice **(s) de Thèse :**

Simon THORPE (CerCo, Toulouse)
Florence REMY-EL BOUSTANI (Cerco, Toulouse)

Jury :

Rapporteurs:
Laurent DEMANY (INCIA, Bordeaux)
Sid KOUIDER (ENS, Paris)

Examinateurs:
Daniel PRESSNITZER (ENS, Paris)
Anne CACLIN (CRNL, Lyon)
Pier-Giorgio ZANONE (UT3 - CerCo, Toulouse)

When noise enlightens: an investigation into the mechanisms of perception and long-term memory for meaningless auditory stimuli.

Jayalakshmi Viswanathan

ACKNOWLEDGEMENTS

It feels surreal to be writing the acknowledgements section of my thesis since I've been dreaming about being a brain scientist since I was 14. Looking back, this would not have been possible without the help and support of several wonderful people I have had the opportunity to know and work with over the years. During the last three years, thanks to all the people mentioned below and more, I have learned a lot and experienced tremendous personal growth. It also feels fitting to be remembering all the adventures I have been on during these years, having spent them trying to understand how memory works.

First and foremost, I would like to thank Dr. Simon Thorpe for having hired me on the M4 project despite intense competition, and for his guidance and supervision over these years. It has been a truly brilliant experience working with him, and I have left every meeting with him feeling inspired, enthusiastic, uplifted and full of ideas for future experiments. He has always encouraged me to look at the bigger picture and try to understand brain function as a whole. As director of the research institute, Simon has also fostered an amazing work environment where students and researchers are encouraged to share ideas and collaborate. As a result, most of my colleagues at CerCo have become some of my closest friends.

I would also like to thank Dr. Florence Rémy for being an amazing co-supervisor and teacher. Under her tutelage, I have grown to be a more effective scientist and science communicator. Whenever I was stuck with an issue while running experiments or analyzing data she provided me with timely and helpful ideas and suggestions, which I believe greatly improved the quality of the research presented here. Having never performed an fMRI experiment before, it was a real pleasure learning the subtleties of fMRI design and analysis from her, topics in which she is an expert. I am also grateful to Florence for giving me the opportunity to train first year Masters (Biosanté) students (UPS) in running experiments; it was a hugely rewarding experience for me.

Next, my gratitude goes to Dr. Nadège Bacon-Macé, the former, and Dr. Karine Bouyer, the current, managers of the M4 project. As a mentor and friend, Nadège helped me successfully integrate into the M4 project both scientifically and administratively. She also helped me settle in Toulouse, all while I was struggling to learn a new language and new experimental techniques. Over the last year and a half, Karine has been an amazing friend and colleague, helping me navigate administrative issues, providing encouragement, technical and organizational support as well as effectively managing the M4 team. She has also helped me by translating several official documents into French, including the "Résumé substantiel" of this thesis.

I would like to very much thank Martin Deudon and Dr. Tim Masquelier who helped me in testing my hypothesis regarding scrambling sounds using the STDP model developed by Tim. Martin built the sound to spike convertor and tested the sounds from my experiment on Tim's model. In addition, I would also like to thank Dr. Trevor Agus for providing me with several MATLAB scripts, for running experiments (which I used to learn MATLAB).

I would also like to thank the members of my thesis jury, especially Dr. Laurent Demany and Dr. Sid Kouider for agreeing to be the reviewers, as well as Dr. Daniel Pressnitzer, Dr. Anne Caclin and Dr. Pier Giorgio Zanone for being on my examination committee. I would also like to thank everyone for taking time out of their busy schedules to attend my thesis defense, some of them coming from very far away.

With regards to the fMRI experiment, there are a few other people I would like to thank. The fMRI technical team at CerCo, especially Dr. Nathalie Vayssière and Dr. Isabelle Berry provided vital help during the data collection period. They were also incredibly patient, friendly and provided valuable suggestions to improve data collection and analysis. Sophie Muratot, our M4 team's research assistant, as well as Chadlia Karoui and Rémy Faucher, my interns for two months, provided invaluable assistance in experimental set up and data collection of the fMRI experiment and have my gratitude. I would also like to thank all the volunteers who have participated in my experiments and thereby helped me answer the hypotheses I was investigating. For three years, the first year Masters (Biosanté) students have helped me in collecting data for various experiments for which I extend my heartfelt thanks to all of them.

All my delightful colleagues in the M4 team and CerCo have incredibly enriched my life over the last three years. Special thanks go to my office buddies, Christelle Larzabal, Danae Remon and Sophie Muratot, who have filled my work days at CerCo with science, friendship and laughter. It was also a real pleasure working with close friend and colleague Marina Inyutina to organize CerCo's Journal Club for two years. Various other researchers, post-docs, fellow doctoral students, masters' students, RAs and interns at CerCo have truly enriched my days with discussions about science, life, philosophy and language and I am very thankful to each and every one of them.

Over the years of working in Neuroscience, I have been fortunate to make amazing friends from all over the world, friends who have supported me in various ways and have been more like family than friends, and I would like to thank them.

Last, but definitely not least, I would like to thank my parents from the bottom of my heart, for always encouraging me to follow my dreams, even when it made life downright difficult for them to do so. I am at a loss for words to describe how grateful I am to them; for always teaching me to do the right thing even when it is difficult, for inspiring me to be a better person every day and for teaching me to be curious about the world. I would not be standing here without their love and support.

ABSTRACT (ENGLISH):

Since the dawn of civilization, the importance of memory in thought, creativity, knowledge and decision making has been recognized. Despite a rich history of philosophical and experimental enquiry into the topic, several questions regarding the mechanisms and neural correlates of memory remain unanswered. Since Guttman and Julesz demonstrated in 1963 that humans are able to detect acoustic features in Gaussian noise, numerous studies have investigated the properties of this capability. In an elegant experimental paradigm, researchers recently used repeating noise segments [*cyclic noises (CNs)*, presenting a segment of noise several times back to back] to investigate long-term sensory memory (Agus et al., 2010). They asked participants to discriminate CNs from plain noise segments, while implicitly presenting them with a few *target* CNs several times. The results demonstrated long-term memory for such sounds, which have raised several questions regarding how and where acoustic features were stored. In this thesis, the robustness of memory for implicitly learned Gaussian sounds was tested using a similar paradigm as the one used by Agus et al. The robustness of participants' recognition memory was evaluated by presenting them with looped and scrambled (10 or 20-ms bin size) versions of target CNs 4 weeks post-learning. Participants were significantly better at detecting cyclic patterns in intact, looped and scrambled versions of target CNs compared to novel CNs, suggesting that neurons might code for very small bits of acoustic information. Additionally, results from this experiment were compared with predictions from a biologically inspired computational spike-time dependent plasticity model *(STDP)* (Masquelier et al., 2008). The compatibility of behavioral and computational results suggests that STDP could possibly be the cellular mechanism implied in storing sensory information. Next, the spatial correlates of memory for implicitly learned Gaussian sounds were explored. Specifically, the role of subcortical areas in storing auditory patterns was investigated in an fMRI setting. Using the same paradigm as in the previous experiment, participants performed the testing session during fMRI scanning. Implicit memory for target CNs (better discrimination than new CNs) was demonstrated, irrespective of performance during the learning session. Functional contrasts implicate lower areas of the auditory pathway, specifically the Medial Geniculate body, as well as the hippocampus, in this long-term memory. Lastly, in a novel paradigm, we further explored the mechanisms and resolution limits of memory for Gaussian noise. Participants were presented with CNs (of different durations) in one ear and plain noise in the other ear, and had to localize the CN. Implicitly, some target CNs were presented several times. Implicit and explicit memory was tested 4 weeks later. Although participants lacked conscious memory for target CNs, they were better than chance at localizing target 10-ms CNs but not novel CNs, even with 8 repeats (80 ms). This whole series of experiments demonstrate surprising findings: 1) the ability to learn and store patterns of acoustic information shorter than 10 ms; 2) this *memory* is sub-cortical, occurring along the early auditory pathway, in regions implicated in *perception* of sounds; and 3) stored acoustic features are not accessible to consciousness.

Key words: Long-term memory, resolution of auditory representations, sensory learning, fMRI, implicit memory, explicit memory, Gaussian sounds.

RESUME (FRANÇAIS):

Depuis l'aube de la civilisation, l'importance de la mémoire pour la pensée, la créativité, la connaissance et la prise de décision est reconnue. Malgré de nombreuses études expérimentales et philosophiques sur le sujet, plusieurs questions sur les mécanismes neuronaux impliqués dans la mémoire restent sans réponse. Depuis que Guttman & Julesz ont montré en 1963 l'aptitude humaine à discriminer les caractéristiques acoustiques de bruits Gaussiens, beaucoup d'études se sont intéressées aux propriétés de cette aptitude. Grâce à un paradigme expérimental judicieux, les mécanismes de la mémoire sensorielle à long terme de bruits Gaussiens ont récemment été étudiés en utilisant la répétition de segments de bruit présentés en continu, ou bruits cycliques (CNs) (Agus et al., 2010). Ces chercheurs ont demandé à des sujets de discriminer des CNs parmi d'autres bruits purement aléatoires, certains CNs *cibles* étant présentés plusieurs fois à l'insu des sujets. Une mémorisation à long terme de ces CNs cibles a été démontrée, soulevant ainsi d'autres questions sur les mécanismes mnésiques sous-jacents. Dans cette thèse, la robustesse de cette mémorisation des bruits Gaussiens appris de façon implicite a été testée à l'aide d'un paradigme inspiré de celui d'Agus et al. Pour ce faire, nous avons testé la reconnaissance implicite à long terme (1 mois) de CNs cibles ayant subi une transformation acoustique : soit le début du son était décalé aléatoirement et le son enroulé sur lui-même (CNs « looped »), soit le son était brouillé par découpage en segments courts (10 ou 20 ms) présentés dans un ordre aléatoire (CNs « scrambled »). Les sujets ont montré une meilleure discrimination des bruits cycliques pour les CNs cibles (par rapport aux nouveaux CNs), qu'ils soient présentés intacts, « looped » ou « scrambled ». Ceci suggère que de très courts segments de bruit peuvent être stockés en mémoire à long terme. Nos résultats sont en accord avec les performances de reconnaissance prédites par un modèle d'apprentissage « Spike-time dependent plasticity » (STDP) (Masquelier et al., 2008), faisant de STDP un mécanisme cellulaire plausible pour le stockage implicite d'informations sensorielles. Par la suite, nous avons recherché les structures impliquées dans cette reconnaissance à long terme. Le rôle des structures sous–corticales a été étudié par IRM fonctionnelle. Nous avons montré une trace mnésique des CNs cibles entendus 1 mois auparavant, indépendante de la performance sur la tâche d'apprentissage implicite (discrimination des bruits cycliques). Cette trace mnésique à long terme pourrait impliquer les premiers relais de la voie auditive, en particulier le corps genouillé médian, ainsi que l'hippocampe. Enfin, à l'aide d'un paradigme novateur et plus risqué, nous avons exploré les limites de cette mémoire auditive des bruits Gaussiens. Nous avons présenté aux sujets des CNs cibles de différentes durées dans une oreille, des bruits purement aléatoires étant présentés dans l'autre oreille. Les sujets devaient localiser le bruit cyclique. Un mois après, les sujets ont montré une reconnaissance implicite de CNs cibles aussi brefs que 80 ms, cette reconnaissance s'améliorant avec le nombre de présentations du CN cible lors de l'apprentissage. Aucune reconnaissance explicite des CNs cibles n'a été observée. L'ensemble de ces résultats démontre de façon surprenante : 1) la capacité d'apprendre et de conserver en mémoire des segments de bruit aussi courts que 10 ms, 2) une trace mnésique sous-corticale de ces bruits, dans les premiers relais auditifs impliqués dans la perception des sons, 3) un accès non conscient à cette trace mnésique. Mots-clés : Mémoire à long terme, résolution des représentations auditives, apprentissage sensoriel, IRMf, mémoire implicite, mémoire explicite, sons gaussiens.

FINANCIAL ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement number 323711—Project M4. I would also like to acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-12-BSH2-0010—Project ELMA.

CONTENTS

IV. Exploring the mechanisms and resolution limits of memory for Gaussian sounds

LIST OF FIGURES

I. Introduction

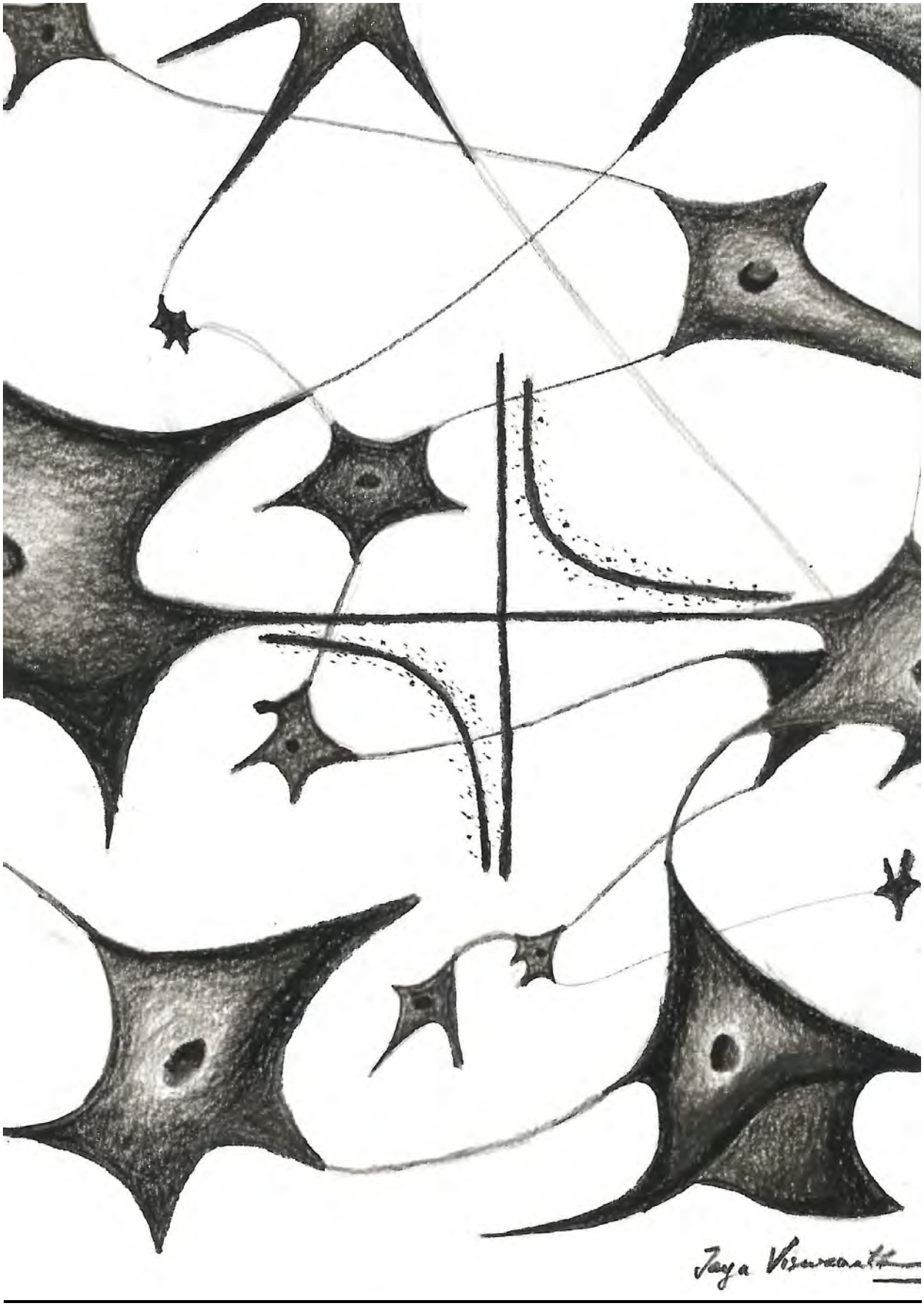
II. Robustness of memory for implicitly learned Gaussian sounds

III. Spatial correlates of memory for implicitly learned Gaussian sounds

IV. Exploring the mechanisms and resolution limits of memory for Gaussian sounds

LIST OF TABLES

LIST OF EQUATIONS


Jaya Visvanath

I. INTRODUCTION

1.1 What is memory and why do we study it?

A) Historical and philosophical investigations into memory

Since the dawn of civilization, the importance of memory in everyday life has been recognized and revered. Sarasvati, the Indian goddess responsible for learning, knowledge and the creative arts is mentioned in the Rigveda, the oldest extant book in an Indo-European language, written between 1700 and 1100 BC (Oberlies, 1998). Ancient Indians were not alone in assuming that memory was the source of creativity: ancient Greeks considered memory to be personified by the Titaness Mnemosyne, daughter of Gaia (earth) and Uranus (sky). Mnemosyne in turn was believed to be the mother of the 9 muses who are the sources of all inspiration for literature, sciences and the arts (Hesiod, 700 BC).

The philosophical inquiry into memory and speculation into its mechanisms began with Plato and Aristotle who theorized memory to be the result of impressions 'stamped' into a mental wax like substance (Aristotle, 350 BC; Plato, 369 BC). Interestingly, Aristotle distinguished memory (of the past) from sense-perception (of the present) besides positing that any animal with the capacity to perceive the passage of time has memory. The next major theory of memory was put forward by St. Augustine between

397 and 400 AD (Manning et al., 2013). He introduced the concept of mental time travel, suggesting that "The time present of things past is memory; the time present of things present is direct experience; the time present of things future is expectation". In retrospect, this idea was a major advance in memory philosophy in two aspects: one since it blurred the rigid distinction in mental processes assigned to the past, present and future actions (a distinction that had been set up by earlier philosophers), and second, since it encouraged understanding memory in terms of its evolutionary function. In other words, this idea supports the notion that need for prediction in everyday life is the evolutionary drive that led to the development of memory systems. The function of the memory system then is to store biologically relevant aspects of the perceived stimuli to help in future action and behavior.

Enquiry into memory remained largely speculative through the renaissance period with several philosophers, mathematicians and scientists like Descartes, Hobbes and Locke putting forward theories of how memory might function (Nikulin, 2015). This continued until Wilhelm Wundt set up the first experimental psychology lab in 1879 to investigate sensory perception and consciousness (Carpenter, 2005). Soon after, Herman Ebbinghaus pioneered running experiments to understand the characteristics of memory, despite the assertions of philosophers of the time (such as Herbart) that

experimental investigation of higher mental processes was impossible (Roediger, 1985). Ebbinghaus was particularly interested in testing the popular hypothesis of memory mechanisms at that time – the association theory. This theory proposed that during learning and memory, like gravity, similar concepts were attracted to each other and were organized/grouped together in the mind (Ebbinghaus et al., 1913). Interestingly, to test his hypotheses, Ebbinghaus used meaningless stimuli – strings of letters to create a series of homogenous, meaningless sequences such as 'mapesch' or 'fajup' – in order to avoid the effect of several parameters which he felt could not be controlled for experimentally. When using real world stimuli, factors such as subjective variability in emotional and attentional response to the stimulus, pre-existing associations and knowledge of the stimulus, the differences in amount and type of information contained in different stimuli etc. will influence a participant's response. Ebbinghaus called these "a multiplicity of influences which change without regularity and are therefore disturbing". By using non-sense stimuli, Ebbinghaus was able to test several specific hypotheses without worrying too much about these unpredictable influences. To date, controlling for these factors is vital in memory research, which makes investigation of memory mechanisms both challenging and rewarding. He was able to test his own memory for nonsense syllables and found that forgetting is essentially

exponential and that without mechanisms to store information in the long term, most of the information encountered is lost within a couple of days. Other researchers have since confirmed this observation (Murre et al., 2015).

Since Ebbinghaus, several psychologists, physicists, physiologists, biologists, neuroscientists, mathematicians and computational scientists have been working on trying to understand how information from different sensory modalities are perceived and stored in humans and different animal models. Despite this long and rich history of philosophical and scientific inquiry into the topic, several questions regarding the mechanisms and neural correlates of memory remain unanswered.

B) Different phases and types of memory

Over the decades of research into mechanisms of memory, the acquisition, processing and storage of information have been defined in terms of different stages of memory formation. Of all the incoming information that is perceived, some information goes on to be *encoded* or learned. Newly learned information is often fragile and easily modifiable. Once learning is complete and performance has reached a plateau (performance no longer improves), memories are considered stable or *consolidated*. After encoding, stored information can be accessed, irrespective of stability, via retrieval processes.

Based on type of learning, memories can be classified as implicit or explicit. In implicit learning paradigms, participants are not made aware of learning. These kinds of paradigms are used to investigate hypotheses regarding automatic learning, subconscious biases etc. On the contrary, explicit learning paradigms where participants are instructed to learn during the experiment are used to investigate specific hypotheses regarding effortful learning and memory.

Memory has been classified into different categories based on different parameters.

For instance, based on the amount of time that information is stored in the brain, memory can be classified as immediate, short and long-term memory (figure 1.1). As demonstrated by Ebbinghaus and other scientists since, most of the information acquired is rapidly lost from immediate and working memory and is not subsequently consolidated into long-term memory (Murre et al., 2015).While this classification holds

true at the individual level, we also carry epigenetic and evolutionary memories spanning several generations. The first documented instance of evolutionary memory was the observation that new-born bird hatchlings display elaborate 'instinctive' behaviors; they automatically crouch when a hawk passes overhead but do not react to other birds (Tinbergen and Niko, 1953).

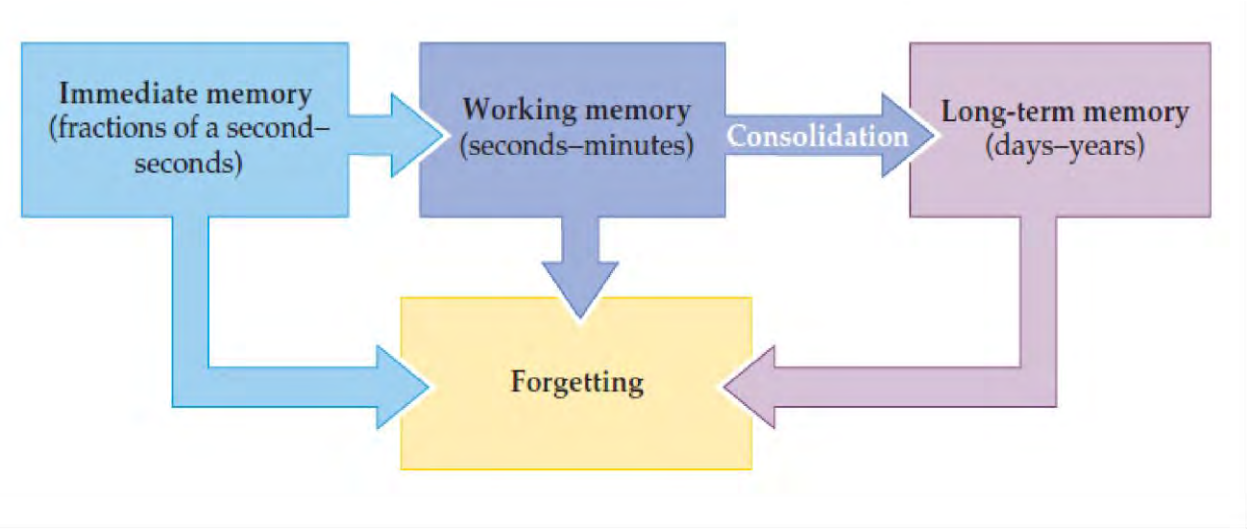


Figure 1.1: The temporal categories of human memory, within an individual (excluding genetic memory). Adapted from (Purves et al., 2008), figure 31.2.

An important point to note here is that the time scales of experiments do not always match the time scales of the phenomenon being investigated. Most experiments investigate phenomenon spanning a few seconds to a few hours. Some memory experiments have investigated events spanning a larger time frame such as days, months or years. However, many of the processes that directly affect parameters being measured in experiments are occurring at a completely different timescale.

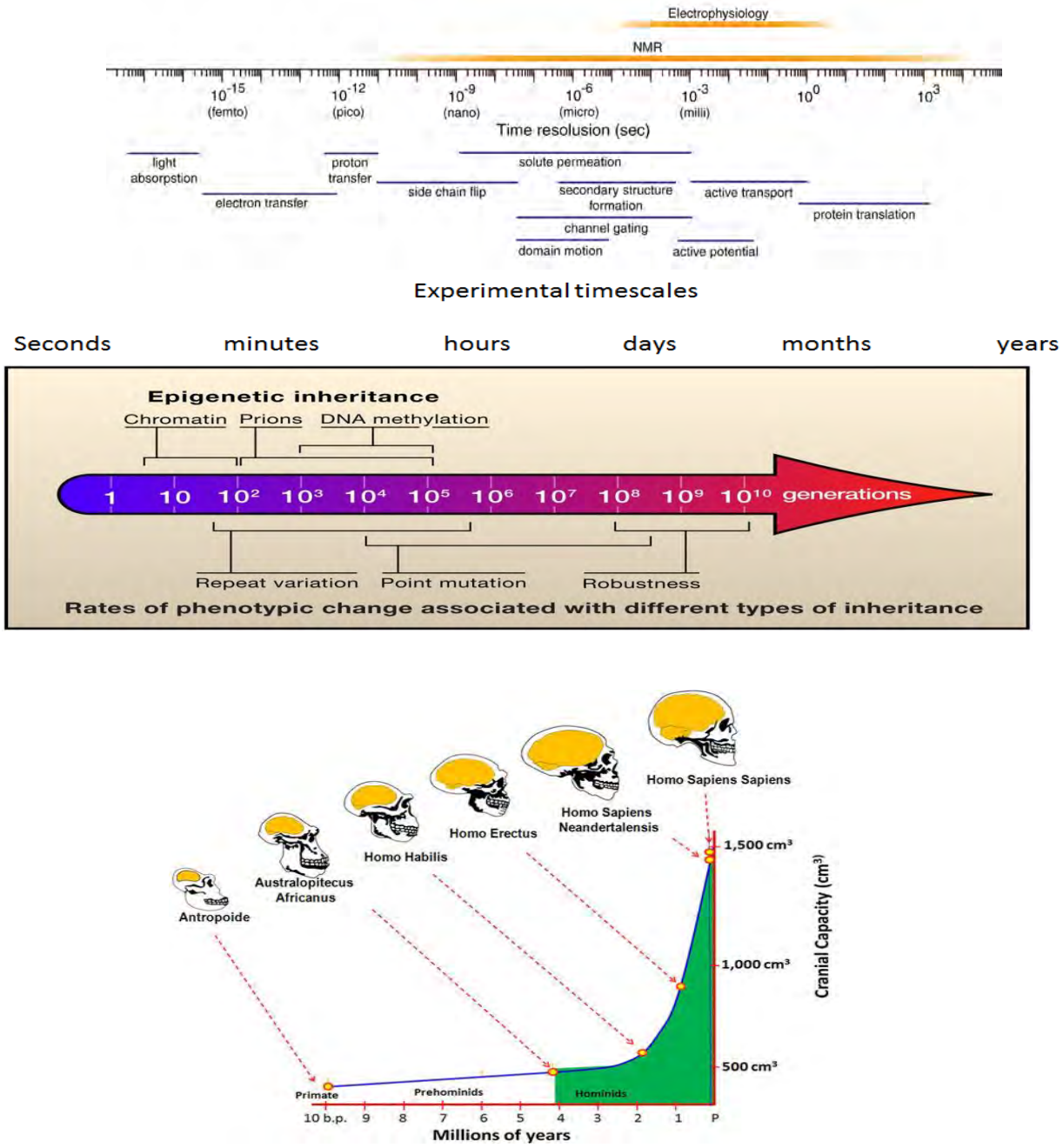


Figure 1.2: Timescales of events at different levels of organization. From the top - panel 1: events at the biophysical level. The timescales of some fundamental atom- or molecule-scale events are also shown. Events include molecular and cellular modifications. Image adapted from (Ode et al., 2012). Panel 2: experimental time scales. While most experiments investigate events ranging from minutes to hours, a few have also investigated events longer durations. Panel 3: events at the epigenetic level spanning several generations. Events involve different levels of DNA modification. Image adapted from (Rando et al., 2007). Panel 4: example of measurable changes (in brain size) observed at the evolutionary time scale scale spanning millions of years. Image adapted from (Díaz, 2013).

It is especially important to keep this in mind when investigating mechanisms of memory since events that impact the encoding and storage of information span several levels of organization.

Examples of events at each time scales are represented in figure 1.2. Molecular scale events such as protein transcription and translation play a role in encoding and perpetuating memory stores. At the experiment time scales, there are different kinds of events which affect memory such as stages of sleep, attentional oscillations and consolidation/stabilization of memories. At the epigenetic level, the changes usually involve temporary (such as methylation) or minor permanent (such as point mutation) modification in the DNA and such information is stored for several generations. Lastly, events at the evolutionary time scale span millions of years and involve massive changes in behavior.

Memories are also classified depending on how accessible stored information is to conscious processing, as shown in figure 1.3. Memories that are accessible to consciousness and can be verbalized, such as episodic memory for specific events in life like a graduation ceremony, are declarative. On the other hand, several memories exist such as motor memories required for swimming or playing the piano, which are not accessible to consciousness. These memories are non-declarative.

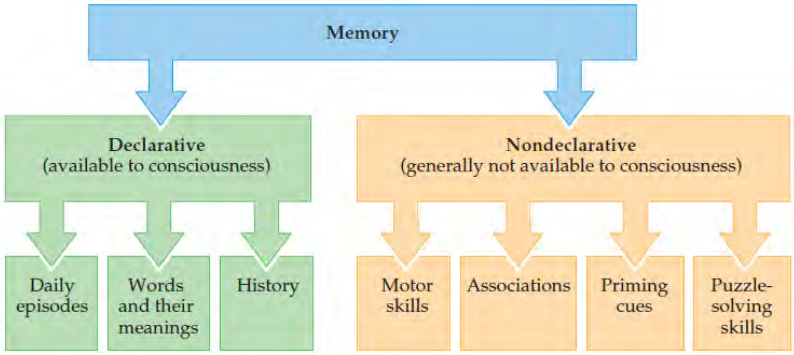


Figure 1.3: Qualitative categories of memory. Declarative memories can be consciously recalled and expressed. Non-declarative of procedural memories are not accessed consciously. Adapted from (Purves et al., 2008), figure 31.1.

Based on the classification of memory systems, it is easy to assume that mechanisms of memory systems are different from each other and boundary conditions between these distinctions are rigid. Is that really the case, though? Are the distinctions between memory systems rigid? Answering this question requires thinking about memory in terms of mental time travel. Mental time travel, defined as the ability to use previously encoded information to create probability based predictions of future events, has been observed in humans (Suddendorf and Corballis, 1997) and animals (Suddendorf and Busby, 2003). In fact, mental time travel has been hypothesized to be crucial for decision making (Boyer, 2008).In a seminal fMRI experiment, the same regions that were involved in recalling past events were also activated when participants were asked to 'pre-experience' a planned future task (Botzung et al., 2008). It is clear from this comparison of retrospective and prospective memory systems that boundary conditions between memory systems are not always rigid. In fact, this makes sense in terms of the evolutionary drives behind development of multiple memory systems – different

memory systems probably evolved to address variations in computational challenges of prediction in different domains. That is, predicting probabilities of outcomes and making decisions in different domains requires information to be processed differently. In such a system, we can then predict mechanistic overlaps based on the type of information being processed as well as the task demands as shown by Botzung and colleagues. Support for the idea that memories are flexibly encoded based on how information will be used for future action/decision, also comes from motor sequence learning studies. Motor sequence learning demonstrates that while a sequence is being learned, movements are slower, more measured and require active attentional stores but with increasing practice, movements become progressively smoother and more automatic. If memory systems are rigid, the same network would be involved in storing this information over time, irrespective of level of training, since the information being learned remains the same. However, this is not observed from a mechanistic point of view since the neural networks involved in storing motor sequence memories change over time, either through real practice or mental imagery (Jackson et al., 2003; Lafleur et al., 2002).That is, a shift in behavior where the action becomes more automatic is accompanied by a change in the mechanism of storing the same kind of information. These findings also highlight the role of 'number of exposures' as an important metric influencing how information is stored in the brain. These studies highlight the need to have better models of memory, perhaps based on task demands and observed behavior rather than the type of information being stored.

C) Perception and memory

Memory and perception have long been considered distinct entities, along with the idea that multiple memory systems exist, as discussed in the previous section. The neural correlates of perception and memory have also traditionally been investigated separately. Memory cannot exist without perception however, and the boundary between sensory perception and sensory memory might not be as rigid as some textbooks (and scientists) would have us believe. The first person to question the existence of rigid memory systems was David Gaffan (Gaffan, 2002). Based on evidence from rodent, monkey and human lesion data, Gaffan noted that when areas involved in storing sensory information (memory) are lesioned, processing of sensory information is also adversely affected.

This created a divide in the scientific community and some scientist set to work trying to prove the existence of memory systems and other scientists set out to try and disprove it. Eventually, electrophysiology and behavioral studies on patients with brain damage (in different areas) were able to show a strong link between impairments in sensory processing and loss of sensory memory. First, amnesic patients with impaired conscious recall (declarative memory) but intact sensory perception demonstrate normal non-declarative memory performance with perceptual learning (Corkin, 1968; Keane et al., 1995). Second, lesions studies in monkey perirhinal cortex, (part of the medial temporal lobe memory system) have shown that sensory learning/memory impairments co-occur with impairments in perception of complex features of stimuli (Bussey et al., 2006; Bussey and Saksida, 2002). In fact, evidence from these studies led

the authors to argue that any memory impairment observed by lesioning the perirhinal cortex are a result of compromised perceptual representations of the stimuli. This led to the development of a "non-modular view where memory and perception depend on the same anatomically distributed representations" (Graham et al., 2010). Understanding sensory processing – how sensory information is extracted from input - is therefore an important aspect of understanding how this information is then stored in the short and long term. Mechanisms of sensory perception and memory might therefore not be that different either. Using sensory stimuli in an experimental setup can therefore be a great tool to understand how features are extracted and subsequently stored by the brain.

1.2 How are memories stored and how are these processes investigated?

A) Mechanisms of information processing and storage

Understanding how information is physically stored within the nervous system is important in order to comprehend how network level activity helps make decisions and guide behavior. Unfortunately, practical constraints limit the ability to investigate a research question spanning multiple levels of organization such as the molecular, cellular, tissue and network levels, primarily because events at each level of organization occur on different timescales (figure. 1.2). However, findings from several experiments investigating a question at various levels of organization when put together help build a comprehensive understanding of the mechanism, much like putting together individual pieces of a jigsaw puzzle to build a whole.

At the cellular level, one way that information can be stored is via alterations in the strength of connections, also called synapses, between neurons. The vast majority of communication between neurons in the nervous system occurs at these specialized junctions. The idea that learning induces changes in synaptic strength was proposed as early as in 1949 by Donald Hebb, a Canadian psychologist who postulated:

”When an axon of cell A is near enough to excite cell B or repeatedly or consistently takes part in firing it, some growth or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.” - (Hebb, 1949)

In other words, Hebb postulated that learning can induce modification of biochemical and electrical properties of a neuron via both pre and post synaptic changes (Sweatt, 2010). Results from several experiments conducted during the interim have confirmed this claim and the short- and long-term modifications at the level of the synapse in response to learning are illustrated in figure 1.4.

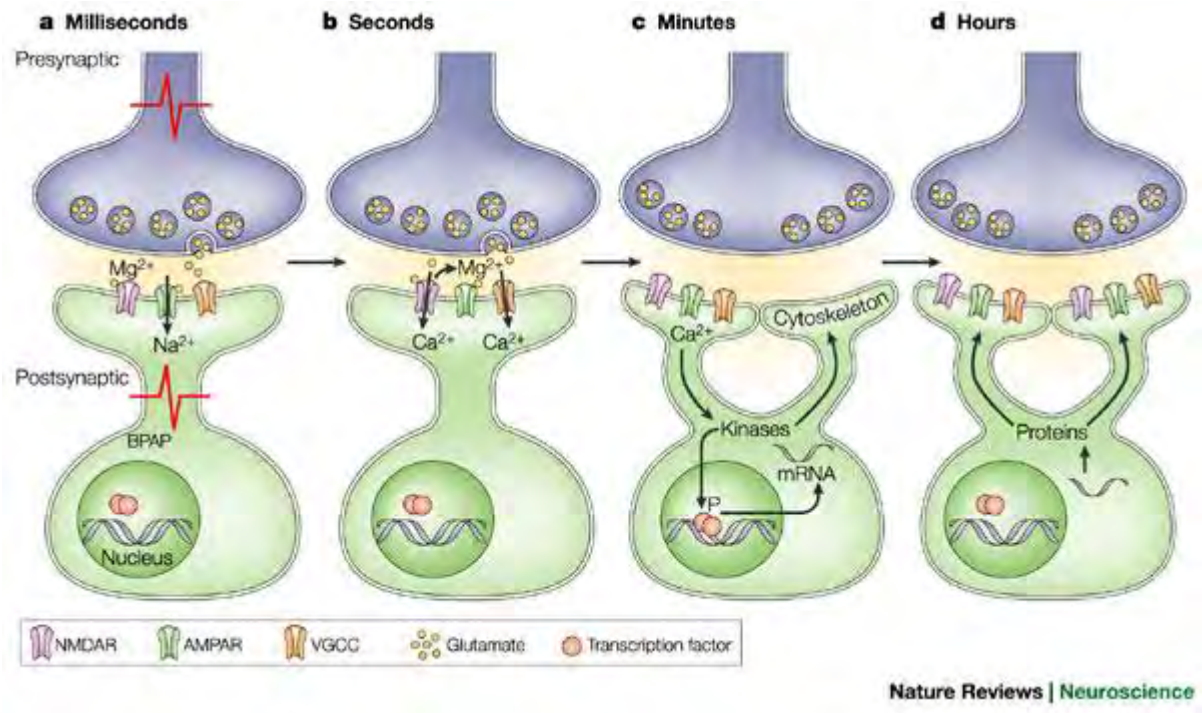
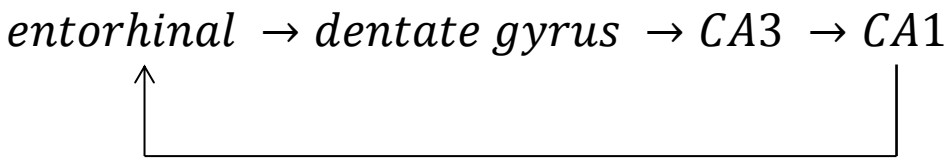


Figure 1.4: A schematic of how information is stored as short and long term changes in the connection between neurons (at the synapse). Spiking activity-dependent release of neurotransmitter (glutamate) from presynaptic neurons leads to the activation of post-synaptic receptors (AMPARs) and to the depolarization of the postsynaptic neuron. Next, depolarization of the postsynaptic neuron leads to local changes in receptor function such as activation of voltage-gated channels. Change in the postsynaptic membrane potential activates protein kinases which, in turn, modulate further gene expression. These substrates contribute to long term local changes at the synapse, such as morphological alteration through cytoskeletal regulation and transcription. Lastly, transcribed mRNA is translated into proteins that are captured by activated synapses and contribute to stabilization of synaptic changes. Image adapted from (Lamprecht and LeDoux, 2004).

At the next level of organization, the discovery of Long Term Potentiation (LTP) spanning multiple synapses helped elucidate mechanisms of learning and memory in the medial temporal lobe (Lømo, 1971a, 1971b, 2003). Synaptic plasticity is the cellular mechanism of storing information in memory and LTP is a form of synaptic plasticity that was observed in the hippocampus and medial temporal lobe (MTL) in response to repeated electrical stimulation. A brief period of high frequency stimulation resulted in a robust but transient[1] (in this case, up to 5 hours post stimulation) increase in the *strength of synaptic connections* between entorhinal (cortical) and dentate gyrus (hippocampal) neurons as well as an increased *likelihood of action potentials.* Three distinct stages were observed in hippocampal neurons– i) immediate LTP was observed right after stimulation, up to 30 minutes post stimulation, mediated by largely unknown mechanisms, ii) early LTP on the other hand lasted between 30 to 120/180 minutes post-stimulation and was mediated by protein kinases and iii) late LTP which was seen after early LTP and up to 300/360 minutes post-stimulation was accompanied by gene expression changes (Sweatt, 2010, chapter 7). This transitional role of the hippocampus evident in LTP response to learning raised several interesting questions regarding the role of medial temporal structures in learning and memory. The main information processing pathway in the hippocampus is the tri-synaptic pathway:



[1] Subsequent research has demonstrated that the duration of LTP effects in the MTL can last a variable amount of time – days, weeks and even months post-stimulation. The duration seems to depend on several factors, including activity patterns in the neurons post-stimulation (Abraham, 2003).

Interestingly, pre and post synaptic changes were observed with LTP at each synapse.

The critical role of medial temporal lobe structures and LTP in memory becomes clearer when considering case studies of patients with damage to these structures. Investigations linking observed behavior to lesions in brain areas are at another level of organization compared to multi-synapse investigations which led to the discovery of LTP discussed above; combining knowledge gleaned from both further extends our understanding of memory processes on the whole. By studying one such patient, HM, researchers were able to establish that these structures are essential for the formation of novel memories, especially consciously recalled declarative episodes. HM suffered from intractable epilepsy and underwent bilateral MTL resection to stop the seizures. This resulted in HM having profound memory defects while retaining normal perception, intelligence and reasoning abilities. His memory defects consisted of a complete inability to form novel declarative memories or recall recently formed memories. However, his ability to form non-declarative memories was intact, as were some early childhood memories. Interestingly, he seemed to be 'stuck' in time for years; he had no recollection of his operation and seemed to think it was a few months before his operation (Corkin et al., 1997; Scoviille and Milner, 1957). Another patient, NA had focal but extensive lesions in the thalamus and anterior temporal lobe due to a fencing accident. Similar to HM, he subsequently had profound anterograde amnesia, or the inability to form novel declarative memories, while his capacity to form non-declarative memories remained normal, as were his memories for events prior to his accident (Squire et al., 1989; Teuber et al., 1968).

These results highlight the importance of MTL structures in forming novel declarative memories. Taking a step back, it can be extrapolated that LTP and the transient increase in connectivity between hippocampal neurons in response to encoding novel memories therefore plays an important role in forming novel declarative memories, but is not involved in retrieval of remote memories that have already been formed. Some kinds of non-declarative memories on the other hand can be formed independently of hippocampal involvement.

By combining results from experiments in different domains, the role of the hippocampus in memory has become clearer. The hippocampus temporarily stores information prior to consolidation into long term memory. In fact, it has been suggested that hippocampal LTP creates a *memory buffer* where newly acquired information is "held" (Rawlins et al., 1985), allowing the hippocampus to perform pattern separation and pattern completion computations. These computations essentially compare newly acquired information to previously held stores and subsequently trigger either consolidation mechanisms to store new information or access previously stored information to result in recall, familiarity or recognition (Yassa and Stark, 2011). Additionally, the hippocampus has been hypothesized to also encode representations of time and therefore be involved in temporal order pattern separation and completion computations during LTP (Rolls, 2013).

After this memory buffer period, however, newly acquired information is consolidated for storage in various cortical areas. The hippocampus also plays a role in mediating this transformation of information into more stable forms in cortical areas and three

mechanistic models have been proposed. The *standard model* proposed that the hippocampus is critical for consolidation of all kinds of information during learning but once memories are consolidated into stable forms, the hippocampus is no longer involved with this information (Squire et al., 1984). However, the discovery that retrieving a memory makes it labile and sensitive to disruption - just like newly acquired information - and that reconsolidation requires protein synthesis challenged this theory (Alberini, 2007; Nader et al., 2000). Another study also showed that the hippocampus is involved in retrieval of remote autobiographical memories (Ryan et al., 2001). These findings led to scientists putting forward the *multiple memory trace theory* which claimed that the hippocampus is involved in episodic learning, retrieval of memories as well as reconsolidation of these retrieved memories (Nadel et al., 2000). This model was challenged by the observation that the hippocampus can also be involved in non-declarative memories as well as patients with hippocampal damage experiencing perceptual difficulties. Extending the role of the hippocampus to encompass such findings, Graham and colleagues proposed the *emergent memory account*, which claims that the hippocampus is recruited to perceptual and memory tasks depending on task complexity and the computations and comparisons required to perform the task (Graham et al., 2010).

Two alternate theories regarding the mechanism of information storage in cortical areas have been proposed. The *distributed* coding theory claims that information is stored in partial form, with information divided and stored separately in different neurons. That is to say, that no single neuron has the "whole" information, rather, neurons each code for an independent bit of information. Recall would then be the result of activating of all

these neurons with separate bits of information and represented as the sum of individual component activations. On the other hand, *sparse* coding theory claims that information is coded for in its simplest and complete form by a few neurons. In other words, a few neurons specialize to respond to certain concepts/information as a whole and reactivating these neuron(s) is sufficient to trigger recall (Olshausen and Field, 2004; Quiroga et al., 2013; Rolls and Treves, 1990). These ideas are illustrated in figure 1.5. Experimental findings in recent years have supported the idea of sparse coding of information which is also a more efficient (in terms of power) mechanism of storage. Main support for this idea comes from the discovery of neurons with invariant responses to a concept – where a 'Jennifer Aniston' neuron was observed to respond exclusively to images of this actor (Quiroga et al., 2005) as well as the discovery of multimodal invariant response to a concept where a neuron was observed to respond to both visual and auditory representations of a concept such as the 'Oprah Winfrey' neuron (Quian Quiroga et al., 2009).

A series of experiments investigating how neurons are recruited to form a memory trace have provided further evidence for sparse coding memory mechanisms (Han et al., 2007; Yiu et al., 2014). In these molecular biology experiments, expression of *Arc* – *activity related cytoskeletal protein* - was used as a marker of neural activity. When a neuron fires, *Arc* is transcribed and stays for 15 minutes in the nucleus before being taken out into the cytoplasm. It is therefore a temporal marker of neuronal activity. By tracking the expression of Arc in neurons, authors investigated specific hypotheses regarding the auditory fear memory trace in mice.

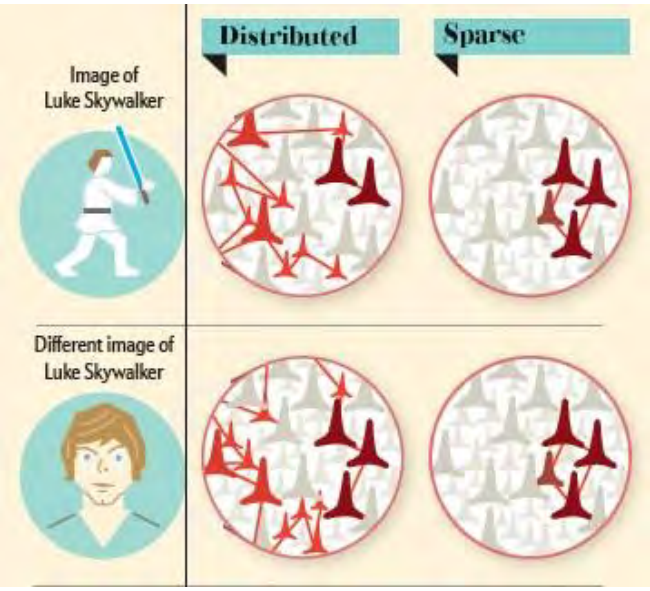


Figure 1.5: This figure illustrates the difference between coding for information in distributed vs sparse networks. An example used here is to code for Luke Skywalker. More neurons are required to store information in distributed networks. It is important to note that with sparse storage 'low' level differences between stimuli – in this case different images representing the same individual, depending on the stage of processing[2], - will be ignored and neurons will only respond to the *concept* of the information stored. Distributed networks on the other hand will detect similarities as well as differences between the stimuli presented to participants. Figure is adapted from (Quiroga et al., 2013).

The authors found that excitability of neurons minutes before and during learning (and not after learning) dictates which neurons are recruited to the memory trace. They were also able to demonstrate that *the same neurons are reactivated during learning and recall.*

[2] Neurons in areas processing low level information regarding objects, such as edge-detector neurons in primary visual cortex, will represent differences in different images of Luke Skywalker; in both sparse and distributed processing. It is important to note that the neurons being discussed here are those that are actually storing representations of the object and sparse and distributed coding theories apply to the final stage of processing when information has to be actually stored, and not intermediate processing stages.

Lastly, they showed that there is inhibition in the network, since size of the memory trace was the same in all control and trial conditions. These findings are in complete agreement with predictions from a sparse coding mechanism of memory.

It is also important to remember that other structures have been implicated in learning and storing information, chief of which is the cerebellum. Neuroimaging studies have demonstrated that the cerebellum plays a role in learning motor sequences, in implicit and explicit memory formation and in working memory (Boyden et al., 2004). The cerebellum has also been implicated in playing a role in supervised learning computations (Doya, 2000). Based on this diverse repertoire of functions and complex morphology of the Purkinje cells that make up the bulk of the cerebellum, it has been hypothesized that the cerebellum plays a role in all functions requiring precise temporal computations (Desmond and Fiez, 1998a). Support for this idea comes from the evidence of cerebellar involvement in auditory working memory where precise computations of fine temporal acoustics is required (Desmond et al., 1997) as well as cerebellar involvement in motor sequence learning where precise computations of fine motor movements is required (Doya, 2000).

The above discussion highlights the complex nature of general memory encoding and retrieval mechanisms, which implies several areas of the brain.

B) Role of attention and sleep on memory

Two factors that have been observed to critically influence how information is stored are attention and sleep.

As early as 1958, Broadbent recognized the importance of attention in perception and subsequent memory and recall, and proposed the *selective filter hypothesis* (Broadbent, 1958). This hypothesis was based primarily on observations from dichotic listening tasks, where participants were able to parcel out and respond to, as well as recall, some information over others in a competing stream. His theory stated that both automatic and voluntary attention acts as filters dictating what information is processed and stored. This theory has been generally accepted and subsequent scientific debates have centered on the site of this influence (Cowan, 1988). That is, how far are stimuli processed before attentional subsampling occurs? The answer to this question seems to depend on the type of information being processed and the task demands. For instance, visual and auditory stimuli seem to be sampled differently, with visual subsampling occurring at the source and auditory subsampling occurring at the level of perception (VanRullen et al., 2014). Prior biases, such as an increased likelihood of detecting your own name, have also been shown to affect how stimuli are processed (Moray, 1959). More concretely, attention has been shown to directly affect behavioral performance measures such as recognition sensitivity, or d' (Moray and O'Brien, 1967; Treisman and Geffen, 1967).

Studies on how sleep influences the encoding and retrieval of memory are also fascinating. Sleep is an interesting subject to study since it consists of multiple stages,

each of which is characterized by specific processes. For instance, the rapid eye movement, or REM, stage of sleep is characterized by muscular paralysis and the electroencephalography (EEG) signatures of brain activity are similar to those of active awake states (15-60 Hz) while slow wave sleep, or SWS, is characterized as the deepest stage of sleep with low frequency (0.5-2 Hz), high amplitude fluctuations in EEG activity. Individual memory traces have been shown to be pruned based on probability estimates of re-encountering the same stimulus (Kim et al., 2014). Sleep has also been shown to selectively strengthen certain individual memory traces over others (Rudoy et al., 2009). How these traces are selected remains unclear but periodic bursts of short duration (1-2 seconds) high frequency (10-12 Hz) oscillations during slow wave sleep, or *sleep spindles*, seem to play a role. Sleep spindles are believed to selectively reactivate, and thereby strengthen memory traces (Marshall et al., 2006; Steriade et al., 1993). In fact, the consolidation of memories has been proposed as one of the main functions of sleep (Sejnowski and Destexhe, 2000). Reactivations during sleep and wakefulness have also been implicated to help strengthen memories in different ways, the summary of which are illustrated in figure 1.6. While running experiments on memory, it is important to note that these two factors will critically influence how and what memories are stored.

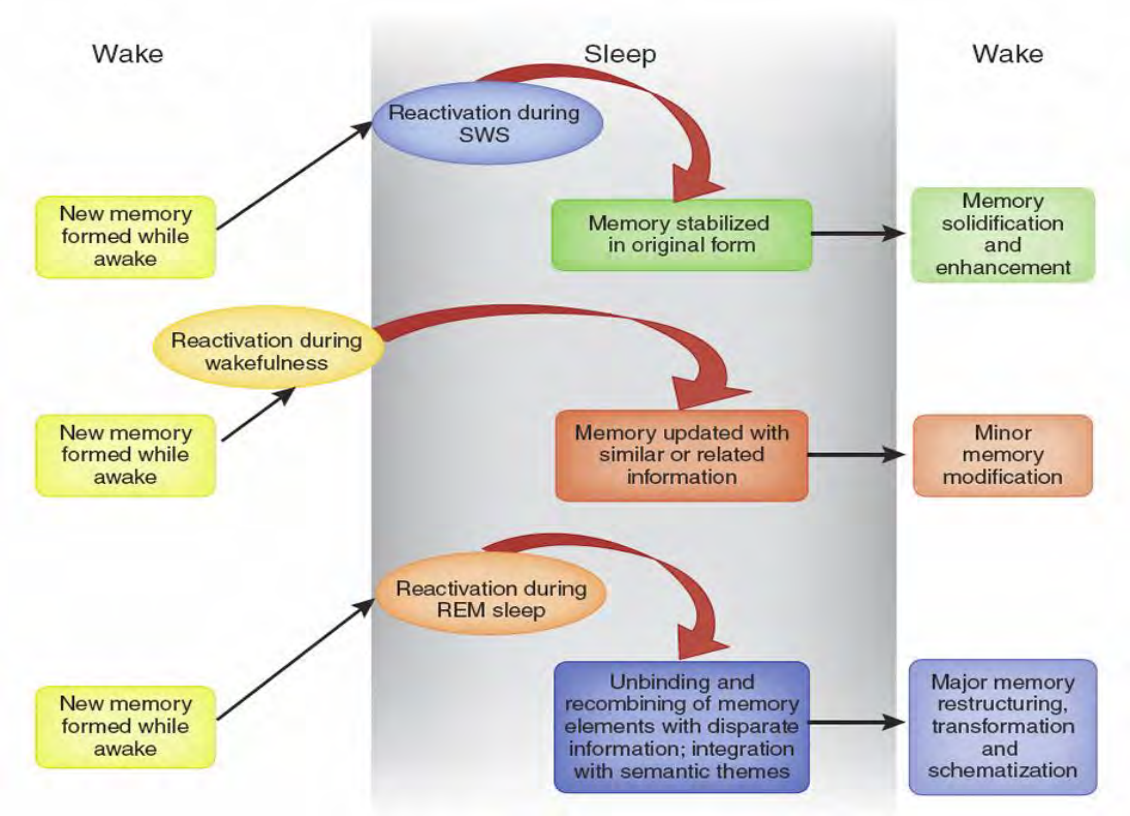


Figure 1.6: A schema of different kinds of reactivations and subsequent consolidation of memories. Top: Reactivation during SWS causes enhancement and stabilization of memory in its original form, leading to memories that are true representations of originally encoded experience. Reactivating memories during SWS by re-presenting a memory cue (such as an odor) present at initial learning leads to memory stabilization. Middle: reactivation during wakefulness causes memory modification and updating, allowing new but related information to be incorporated into the original memory trace. Bottom: reactivation during REM sleep causes substantial memory restructuring and recombination of memory fragments that become isolated in the REM-sleep brain state. Such recombination may lead to insights, creative solutions to problems and memory schematization. Figure from (Payne, 2011).

C) Experimental methods to investigate capacity and mechanisms of memory

Given the complexity in investigating memory systems as highlighted by Ebbinghaus (section 1.1A); three major experimental methodologies have been used to investigate hypotheses regarding mechanisms of memory.

1) Experiments investigating memory for stimuli that were encoded naturally, prior to experimentation.

Experiments testing memory for information that was encoded naturally, also known as experiments using the cross-sectional approach, have investigated truly real-world memories. Results from such studies have yielded more biologically relevant results, but the disadvantage of adopting this methodology is that several parameters cannot be controlled for. An example of such a parameter is the number of times a particular event has been perceived and is consciously recalled, which is naturally variable and thereby confounds measurements of observed behavior. However, experiments using such techniques have shed light on characteristics of memory. In one such study, participants were asked to match names to the faces of their high school classmates up to 15 years post high school graduation. The authors found that performance in this task was very high, at about 90% accuracy, highlighting the role of social, emotional and personal relevance in strength of encoding (Bahrick et al., 1975).

In a very interesting study using slightly modified versions of coins as visual stimuli, the authors were able to show that participants were unable to distinguish distractor

versions of the pennies from the original (Rubin and Kontis, 1983). These stimuli are used with very high frequency and are naturally very strongly encoded. Despite such strong encoding, not all features are stored as demonstrated by this experiment. Therefore, memory is not just a function of repetition, but somehow behaviorally relevant features are encoded. Therefore, these kinds of experiments have led to a deeper understanding of memory functioning under natural conditions.

In our team, as part of her doctoral research, Christelle Larzabal has been investigating memories for TV shows that were aired decades ago and then subsequently taken out of broadcasting circulation. Some of these TV shows were broadcast very few times and others were broadcast several times over a few years. Fortunately, she has access to precise information regarding how many times each of these shows were broadcast, and was able to test older participants memories for different shows they had seen variable number of times several years previously. Her findings suggest that participants are able to form highly detailed and specific memories, sometimes based on 'one-shot' experiences that were strongly linked with episodic memories, despite being dormant for several years and highlight the interest in running such types of experiments.

2) Experiments investigating memory for stimuli that are encoded and then tested some fixed period later in the laboratory.

Experiments using this methodology allow the testing of specific hypotheses regarding memory mechanisms, since parameters that affect learning and memory can be controlled for in laboratory based experiments. These experiments have been conducted

using natural or artificial stimuli. One study using real world stimuli asked participants to watch TV programs in the laboratory context. Participants were later (via surprise memory tests immediately after watching the shows as well as 24 hours later by phone) tested for their recall of advertisements, which were presumably encoded incidentally (Bushman and Bonacci, 2002). Their results indicate that when the content of TV shows were intensely emotional with powerful (violent or sexual) content, participants did not encode advertisements, and were therefore less susceptible to external attentional distractions. Artificial stimuli, such as creating a set of rules to result in artificial grammar compositions, have been used to study the mechanisms and neural correlates of acquiring novel learning rules (Skosnik et al., 2002).

It is important to note that experiments using this methodology can flexibly test several parameters of memory longitudinally. Having participants return to be tested on the same (real or artificial) stimuli several months and years post-learning allows tackling mechanisms of very long-term memory.

3) Experiments investigating memory in patients with brain damage using real world and artificial stimuli.

Most of our knowledge on brain structures involved in memory comes from studying patients with brain injury and impaired memory issues. Unlike dementia, which is a disorder characterized by progressive memory decline, acute brain injury as caused by trauma, stroke, etc., can help isolate the areas critical to certain memory functions as was the case for HM and NA discussed earlier (section 1.2A).

While there are advantages and disadvantages for all of the above methodologies, in this thesis, the focus has been on understanding the mechanisms of memory for meaningless auditory stimuli in healthy participants. The stimuli were encoded and retrieved in a laboratory setting.

1.3 How is sensory information processed and stored?

A) Introduction to sensory memory and models.

In the previous section, the emphasis was on understanding general mechanisms of encoding and storing information. Sensory memory is fascinating to study since it straddles the dual worlds of perception and memory, as discussed in section 1.1C. A major component of sensory perception post-infancy is categorization of percepts. To recognize the form of a face in a photograph, for instance, there has to be a prior 'template' of sorts that tells us that the visual percept in the photograph matches the expected form of a face and therefore, is categorized as such. A completely novel percept can be categorized as such only if no matching template exists in memory. This decision can only be made after comparing the novel percept with each previously stored template. As such, in terms of philosophy as well as brain mechanisms, where perception ends and memory begins is not clear.

From an evolutionary perspective, taste memory is essential for survival. Any problems with ingesting or digesting a certain food can be associatively linked to either aversive or pleasant sensations to discourage or encourage repeated ingestion of the same item. This association must necessarily happen hours after ingestion and theoretically proves the existence of purely sensory taste memory (Bermúdez-Rattoni, 2004). The orbitofrontal cortex (Thorpe et al., 1983) and amygdala have also been shown to store

taste information with both negative and positive sensations (O'Doherty et al., 2001). Interestingly, purely sensory memory has also been observed in other modalities where storing such information is not critical for evolutionary survival. Memory for purely tactile information, such as the ability to consciously distinguish textures, has long been observed corresponding to activity in the somatosensory cortex (Zhou and Fuster, 1996). The same holds true for olfactory memory: the ability to distinguish labelled and unlabeled odors has been linked to activity in the temporal and orbitofrontal cortices (Jonesgotman and Zatorre, 1993). More recently, memory for purely sensory information has been also been demonstrated for visual, purely sensory 'meaningless' information (Gold et al., 2014) and for auditory purely sensory 'meaningless' information (Agus et al., 2010; Guttman, N., and Julesz, 1963). Therefore, the ability to store purely sensory information seems to be generalizable across modalities. But how are these memories stored?

At the cellular level, scientists have proposed a mechanism by which inhibitory GABAergic[3] neurons control the excitability and receptive field sizes of cortical neurons storing somatosensory (Dykes, 1997) and olfactory (Kaba and Nakanishi, 1995) information encoding. The authors thus suggest that this mechanism can be generalized to explain how perception and memory work hand in hand in the same anatomic locations for purely sensory memory in all modalities. In this mechanism, once sensory information is stored in a neuron, inhibitory GABA neurons reduce the excitability of

[3] GABAergic neurons are generally small inhibitory neurons that function via the neurotransmitter Gamma Amino Butyric Acid. When an action potential arrives at a synapse where the presynaptic neuron is GABAergic, GABA is released into the synapse that (unlike excitatory neurons) induce the post synaptic receptors to be inactivated preventing the post-synaptic neuron from firing an action potential.

this neuron to other input so the neuron now responds more selectively to the stored sensory percept.

A caveat of Broadbent's attentional filter theory (discussed in section 1.2B) was the sensory buffer idea. The attentional filter theory held that incoming information is subsampled by attention which gave birth to the idea that all incoming sensory information is held in a 'buffer', a temporary state, before attentional subsampling occurs. The idea then followed that such high resolution, meaningless sensory information is only stored in short-term memory and only meaningful, lower resolution sensory information that is associated with some semantic label can be stored in long term memory (Broadbent, 1958; Lachter et al., 2004).

For several decades after Broadbent proposed his theory, therefore, high resolution sensory memory was thought to be characterized by 4 features – a) it forms independently of attention since it is held in memory prior to attentional subsampling, b) it is modality specific, c) it has fine resolution and d) it has a short retention time. Sensory memory is thus different from categorical memory which is held in long term memory.

Cumulative evidence against this model of sensory memory was discussed by Winkler and Cowan (Winkler and Cowan, 2005) in light of results from auditory memory reactivation studies. Results from several studies using the mismatch negativity paradigm suggest that longer lasting memory exist for acoustic regularities that are associated with "anchoring" features of a stimulus. Violations of expectations of these regularities then result in the mismatch negativity related evoked potential. In one such

study, authors used trains of tone patterns and demonstrated that participants can hold sensory information for longer than 30s, since mismatch negativity evoked potentials were observed for violations of expectation over this large time window (Winkler et al., 2002).

What these anchoring features might be, or how these regularities might be stored, is not clear but Winkler and Cowan argued for a need for better models to explain sensory memory (Winkler and Cowan, 2005). In fact, while the immense capacity to store sensory information in immediate memory is recognized, few studies have investigated the existence of purely sensory information in long term memory (Purves et al., 2008), despite empirical evidence to the contrary. For example, there is empirical evidence that sensory long term memory for voice features exist, based on our capacity to use purely acoustic feature information to make identity judgements (Craik and Kirsner, 1974).

As argued by the studies discussed above, there is a need for a better model to explain mechanisms of sensory memory.

In 1963, Guttman and Julesz were able to demonstrate that participants store meaningless, purely sensory information in working memory. By 'freezing' a segment of auditory white noise and playing it several times back to back, they created a *cyclic noise (CN)*. They were able to show that within a few presentations, participants were able to perceive cyclicity by detecting some information as features that re-occur rhythmically. Participants identified these features, or brief auditory percepts, as "clunks" and "whooshes". This paradigm, also called the frozen-noise paradigm (Guttman, N., and Julesz, 1963), is the inspiration for the stimuli used in all the

experiments described in this thesis investigating sensory memory in the auditory modality. There are several advantages in using these stimuli to investigate sensory memory, as highlighted in the following paragraphs.

One reason to study sensory memory using meaningless stimuli is that these stimuli straddle the gap between declarative and non-declarative memory systems and therefore conscious and non-conscious processing. While global features that repeat are consciously detected and responded to, participants are unable to have access to individual feature information and cannot consciously recall individual exemplars they have previously heard.

Another reason to use these stimuli goes back to the stimuli used by Ebbinghaus, discussed in section 1.1A (Ebbinghaus et al., 1913). Ebbinghaus wanted to use meaningless stimuli to eliminate biases due to emotional responses, pre-existing associations and biases as well as different amount of information contained in different stimuli. In order to do so, he used strings of nonsense letters. However, these stimuli were not perfect since some strings contained partial phonemes that made them more memorable than others. Gaussian sounds, on the other hand, do not show any such identifiable fluctuations in amplitude to make one segment stand out compared to others. It is also possible to generate large, non-overlapping stimuli sets.

Lastly, since sensory memory forms the basis of all of our interactions with the world, the formation of sensory memories, and their subsequent meaningful associations to relevant objects in our environment, is a building block of learning during infancy when all stimuli are first meaningless. Learning during infancy is purely driven by repetition

of stimuli and the probability of re-encountering said stimuli. Evidence for such early statistics/probability-based learning comes from language learning studies in infants. In one study, 16.5-month old infants were shown to rapidly pick up acoustic regularities in phoneme sequences and look for subsequent violations of the 'new' rule when encountering novel stimuli (Chambers et al., 2003). Results from another study showed that two minutes of exposure were sufficient for 8-month old infants to be able to segment words from fluent speech, based on computation of purely statistical/meaningless information present in the language (Saffran et al., 1996). Further, this learning mechanism was shown to be domain non-specific, with infants similarly learning visual statistical information (Kirkham et al., 2002). These studies highlight the role of probability of co-occurrence in statistical learning, and association of a meaning to this co-occurrence can occur at a later stage. Understanding memory for meaningless stimuli could therefore expand our understanding of statistical learning during infancy. It is also clear that humans retain this ability to learn sensory information purely based on repetition as adults in real world situations - for instance, adults are able to acquire skills such as differentiating bird calls based on purely acoustic information present in the stimuli. Therefore, understanding mechanisms of sensory memory in adulthood could shed light on the mechanisms of learning in infancy. Further extending this idea, even animals have been shown to be able to learn meaningless information either via associations and/or repetition methods. In one study, a border collie was trained over a period of three years to learn the names of 1022 objects, purely based on repetition of the association between the object and the name (Pilley and Reid, 2011). Another study showed that even wolves, who usually do not

respond to human social cues the way dogs do, can with extensive training learn to respond to human gestures that are essentially meaningless for them (Udell et al., 2008). Interestingly, the strategy of identifying novel objects using a process of elimination of already known objects, was also demonstrated in a border collie (Kaminski et al., 2004). This process of elimination, called 'fast mapping' has been identified in human children during speech acquisition. It gives further evidence that mechanisms of statistical learning have been preserved across species due to the importance of probability-based estimates for survival. The discovery of this similarity of learning mechanisms between dogs and children has inspired another doctoral candidate in our team, Danae Remon, to systematically investigate this question.

To conclude, there are several hypotheses of interest to investigate regarding the mechanisms and neural correlates of sensory memory. Due to the universal nature of statistics based memory that has been observed in different sensory modalities, in infants and even animals, understanding how such information is processed and stored would further our understanding of memory mechanisms in a fundamental way.

B) Sensory coding and STDP

Understanding the sensory code involves solving the problem of how sensory information is perceived, transmitted, decoded and stored by a system. This problem is clearly not a straightforward one. So, what do we know regarding how sensory information is coded in the brain? What is the neural code and how is sensory information transmitted and stored?

In the auditory modality, the neural code for perception and memory must necessarily transmit high resolution temporal information in order to perform the precision computations required for tasks, such as sound localization and pitch discrimination. Such a neural code would have to balance this task requirement with a need to minimize energy costs and maximize efficiency of computations (Smith and Lewicki, 2006). Tonotopy, or the hierarchical organization of frequency responsive neurons, is preserved throughout the auditory processing pathway (described in section 1.4A) and it seems that preserving the temporal (time coding) and spatial (place coding) patterns of stimulation at the auditory periphery is important in audition (Evans, 1978).

In recent years, advances in machine learning algorithms have led to the creation of artificial systems capable of performing tasks with levels of efficiency similar to or better than humans. A good example of such an artificial system is a deep neural network with the capacity to classify 1.2 million, high resolution images with very high accuracy levels, into 1000 classes (Krizhevsky et al., 2012). After extensive training, the

classifier had a record low error rate of 18.2% for the top 5 answers provided for each test image, examples of which are illustrated in figure 1.7.

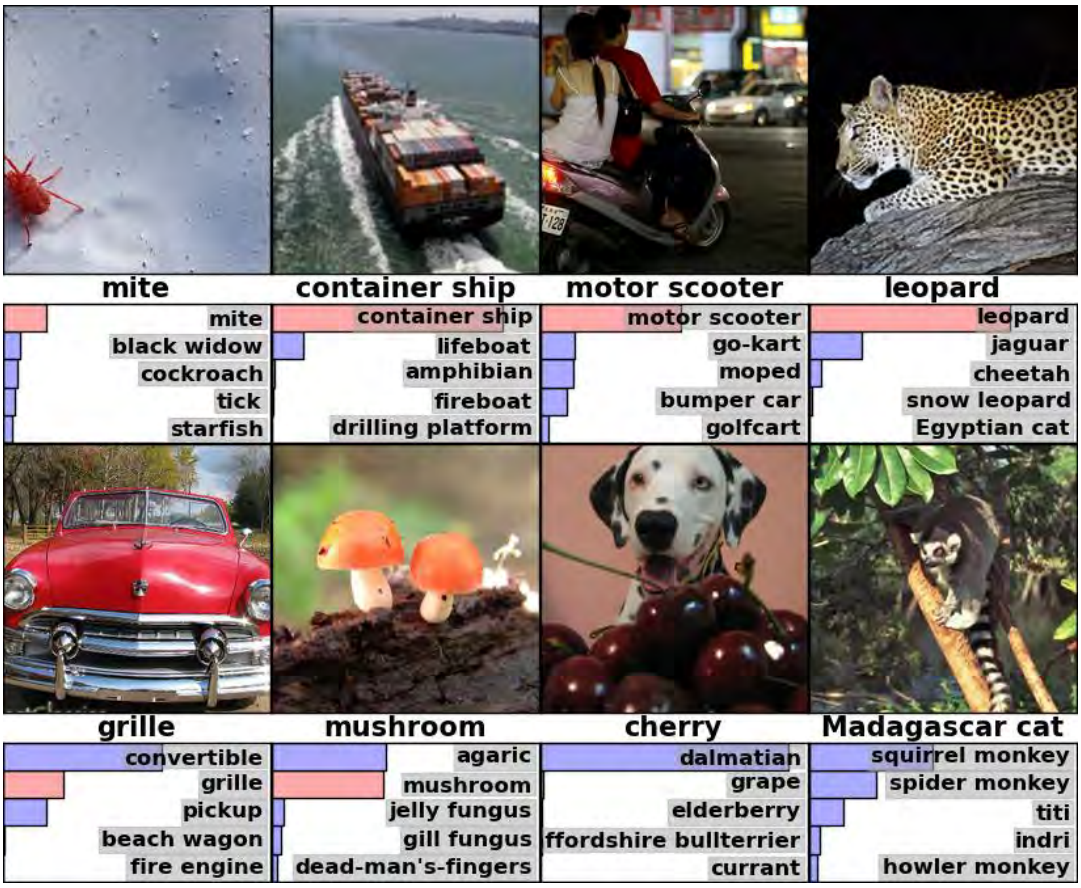


Figure 1.7: Results obtained by the convolutional neural network classifying test images post training. Image adapted from (Krizhevsky et al., 2012). The five labels under the images are the labels considered most probable by the model. The correct label is written under each image, and the probability assigned to the correct label is also shown with a red bar.

The performance of such an artificial system capable of performing at levels of human or better levels solves a problem from a computational point of view. However, it is not known if similar computational mechanisms are used by real biological systems to perform similar task. In other words, how biologically relevant is this computational

mechanism? Does human image classification work the same way as deep convolutional networks?

One way to answer these questions is to build computational systems that are based on what we already know about biological systems, and then compare performance between biologically inspired and purely computational mechanisms for a particular neural computation application.

One of the fundamental mechanisms by which unsupervised learning is achieved in biological systems is Spike-Time Dependent Plasticity (STDP), a form of Hebbian synaptic plasticity. Essentially, STDP (which has been observed both in vitro and in vivo in several brain regions in different animals) regulates synaptic strength based on the timing of events. Spike events that occur closer together in time strengthen a synapse and events that occur further apart in time weaken it. That is, the connection between the pre-synaptic neuron and post-synaptic neuron is strengthened when input spikes and output spikes are tightly coupled in time. On the contrary, synapses are weakened when the input spikes occur just after an output spike. In vitro, this window of causative coupling due to STDP has been shown to be 5-20 ms before post-synaptic spike[4] (Zucker et al., 1991). The observed change in synaptic strength (measured as excitatory post-synaptic potentials) with changing the spike timing is shown in figure 1.8A. This form of plasticity is clearly advantageous for learning since only *causative* connections are preferentially reinforced.

[4] Correspondingly, when pre-synaptic spikes occur 5-20 ms after the post-synaptic spike, the synapse is weakened.

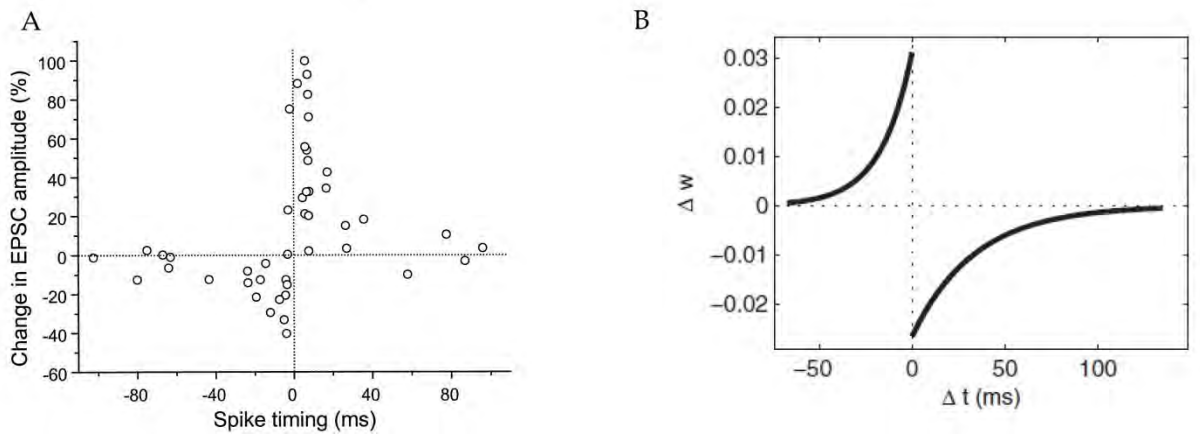


Figure 1.8: A) The observed change in synaptic strength, measured as a change in excitatory post-synaptic current (EPSC) with changing time intervals between the onset of the input compared to output spikes. Change synaptic strength is higher when the difference in spike time is close to zero. Data from glutamatergic neurons from embryonic rat hippocampal cultures; image adapted from (Zucker et al., 1991). B) The STDP learning rule used by a model neuron presented with randomly repeating patterns in noise. Left side represents long-term potentiation and the right side represents long-term depression. Image adapted from (Masquelier et al., 2008).

In recent years, STDP has inspired computational neuroscientists and engineers to find biologically inspired solutions to several computational tasks. One team of researchers implemented STDP with a nano-device and created a memristive[5] device capable of performing efficient unsupervised learning and demonstrating plasticity. Bichler and colleagues created a silicon spiking retina with STDP capabilities and used this to build a spiking camera with dynamic vision to track traffic (Bichler et al., 2012). This camera is much more energy and cost efficient compared to current video based traffic cameras (figure 1.9).

[5]Memristive devices are electrical resistance devices capable of storing and processing information.

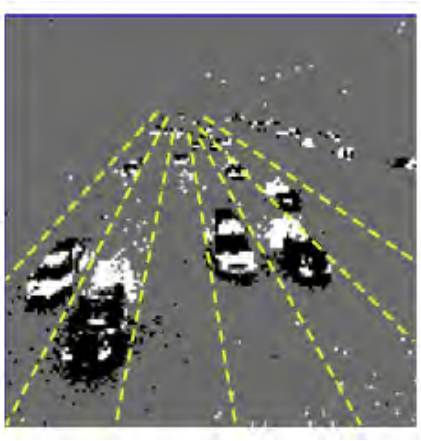


Figure 1.9: Sample image from STDP-based vision sensor used to record traffic on a real highway in Pasadena, California. Image adapted from (Bichler et al., 2012).

The most relevant application of STDP based computing system to research in sensory memory has been the use of STDP models to learn repeating patterns in continuous noise. How patterns are recognized and learned with millisecond precision, when randomly embedded in noise, is a difficult computational challenge to solve. Using an STDP model of a leaky integrate and fire neuron with random Poisson activity, Masquelier and colleagues were able to solve this challenge (Masquelier et al., 2008, 2009, 2016). The STDP learning rule used here is shown in figure 1.8B. The neuron had 2000 afferents with variable instantaneous firing rates at baseline. Continuous spike trains were fed to this neuron capable of detecting spike coincidences, and an arbitrary pattern was randomly repeated in the input stream. The target pattern consisted of 50 ms of spiking activity copy-pasted at random intervals in a subset of the afferents (figure 1.10). The neuron was able to specialize and respond with 100% selectivity (0 false alarms) to the target pattern within few tens of presentations, demonstrating fast, unsupervised learning. After learning of the target pattern, a small fraction of the

synapses had become selective to the pattern (383/2000 afferents) and the rest were completely silent (figure 1.11). Interestingly, while chance determined which part of the 50 ms target pattern the neuron learned, the first spike was observed as early as 4 ms after target pattern onset, suggesting that really small features of repeating patterns were detected.

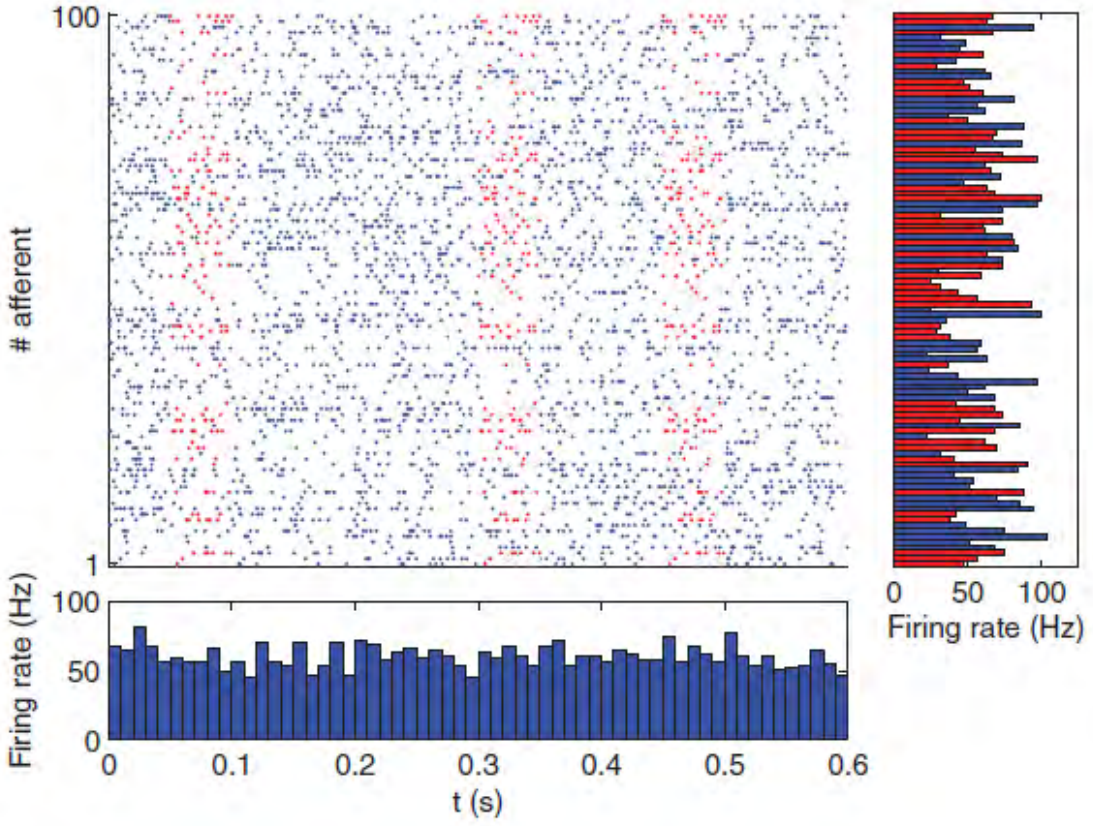


Figure 1.10: Spatio-temporal spike pattern. A repeating 50 ms long pattern (shown in red) randomly pasted into a continuous spike train. The bottom panel plots the population-averaged firing rates over 10 ms time bins, and demonstrates that nothing characterizes the periods when the pattern is present. The right panel plots the individual firing rates averaged over the whole period. Neurons involved in the pattern are shown in red. Again, nothing characterizes them in terms of firing rates. Detecting the pattern thus requires taking the spike times into account. This is the input to the neuron with STDP capability. Adapted from (Masquelier et al., 2008)

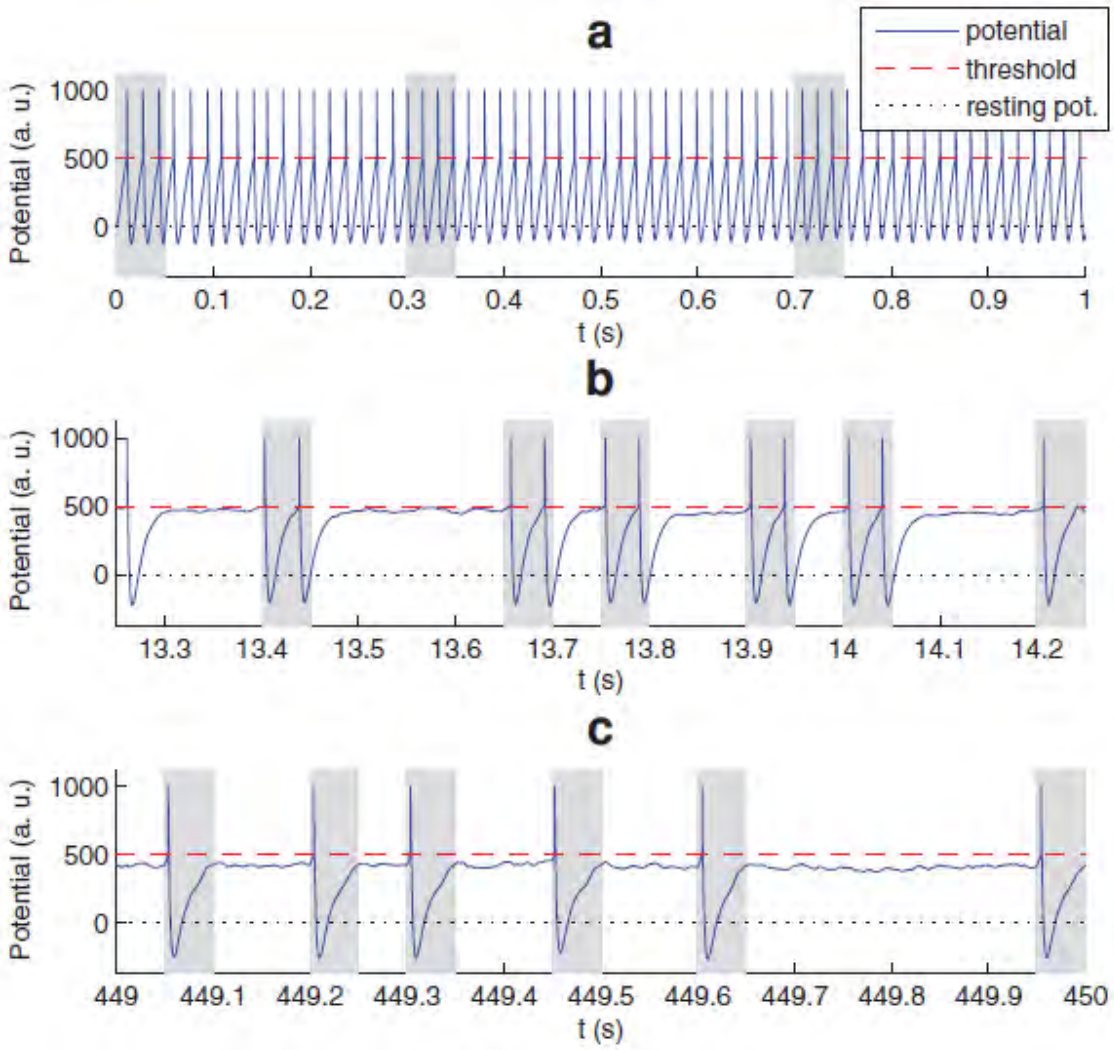


Figure 1.11: Evolution of output spiking activity of the STDP neuron fed with repeating patterns in noise with progressive exposure to the repeating segment. (a) At the beginning of the simulation the neuron is non-selective because the synaptic weights are all equal. It thus fires periodically, both inside and outside the pattern. (b) After about 70 pattern presentations selectivity to the pattern is emerging: gradually the neuron almost stops discharging outside the pattern (no false alarms), while it does discharge most of the time when the pattern is present (high hit rate), here even twice (c) End of the simulation. Postsynaptic spike latency is about 4 ms. Hit rate is 99.1% with no false alarms (estimated on the last 150 s). Figure adapted from (Masquelier et al., 2008).

Interestingly, cortical neurons with firing rates of 25 Hz or lower have been shown to function as coincidence detectors (König et al., 1996). Since any coincidence detecting neurons with STDP can learn patterns based on repetition alone, these types of neurons might be involved in perceiving cyclicity present in meaningless auditory noise. While Konig and colleagues studied cortical neurons, any low firing rate, coincidence detecting neuron, either cortical or sub-cortical, could theoretically perform such statistical learning function in any sensory modality.

Putting all these together, a possible mechanism of explaining how random patterns are learned in noise using STDP comes from stochastic resonance. Stochastic resonance is a phenomenon, where *optimal*[6] noise can enhance the periodicity of a weak signal causing the signal to rise above the threshold for detection (Wiesenfeld and Moss, 1995). This phenomenon was demonstrated in the somatosensory system (in anesthetized cat): periodic tactile stimuli, which had been optimally enhanced through addition of noise, evoked field potentials (Manjarrez et al., 2003). Similarly, acoustic features (weak signal) of a Gaussian sound may be preferentially enhanced when added with baseline neural activity (optimal noise) for a given individual, resulting in different features being learned. Hypothetically, such features that are detected and cause an action potential on successive presentations of the pattern can trigger a coincidence detecting STDP neuron to learn a random repeated pattern in noise. These ideas raise interesting hypotheses.

[6] The noise has to be optimal since too much noise can mask the weak signal completely and too little noise does not lead to a correlation between the weak signal and detection of events.

C) The M4 project and hypothesis regarding how information is encoded and stored

Based on various ideas discussed here, such as sparse coding of memories (section 1.2A), neurons specializing to detect certain features (section 1.2A), and STDP as the probable mechanism by which neurons learn random repeating patterns in noise (section 1.3B), one of my thesis directors, Dr. Simon Thorpe, set out to investigate a set of 10 rather provocative claims regarding the mechanisms of memory, as part of the European Research Council's M4 project – Memory mechanisms in man and machine. The 10 claims of the project are listed below:

1- Humans can recognize visual and auditory stimuli that they have not experienced for decades.

2- Recognition after very long delays is possible without ever reactivating the memory trace in the intervening period.

3- These very long term memories require an initial memorization phase, during which memory strength increases roughly linearly with the number of presentations.

4- A few tens of presentations can be enough to form a memory that can last a lifetime.

5- Attention-related oscillatory brain activity can help store memories efficiently and rapidly.

6- Storing such very long-term memories involves the creation of highly selective "Grandmother Cells" that only fire if the original training stimulus is experienced again.

7- The neocortex contains large numbers of totally silent cells ("Neocortical Dark Matter") that constitute the long-term memory store.

8- Grandmother Cells can be produced using simple spiking neural network models with Spike- Time Dependent Plasticity (STDP) and competitive inhibitory lateral connections.

9- This selectivity only requires binary synaptic weights that are either "on" or "off", greatly simplifying the problem of maintaining the memory over long periods.

10- Artificial systems using memristor-like devices can implement the same principles, allowing the development of powerful new processing architectures that could replace conventional computing hardware.

The experiments I have conducted as a part of my thesis over the last three years have been aimed at explicitly testing some of these claims. Initially, I was especially interested in tackling the second claim. One of the biggest confounding factors affecting results in any memory experiment is reactivation. Reactivation can be voluntary, with participants explicitly recalling some stimuli they had seen/heard during encoding or involuntary, with some stimuli getting reactivated during sleep but not others. In either case, reactivations in the interval between learning and retrieval will result in variations

in how well memories are maintained over time. This is a source of noise in the measurement of interest – the strength and ability to recall items. However, using Gaussian noise to probe memory mechanisms, it is possible to test the second claim. Since participants do not have conscious access to individual features that are stored, memory for these stimuli cannot be consciously reactivated. While sleep dependent reactivations would still affect memory for individual features that are stored, we can assume that *all* learned features are subject to similar probabilities of such reactivations. Therefore, by using meaningless auditory white noise to test hypotheses regarding sensory memory, I was able to test the second claim.

I was also interested in testing the third, fourth and fifth claims using these stimuli. In order to test the third claim, I have investigated the mechanisms of both the encoding/memorization phase and not just the mechanisms of memory. Quality of encoding depends on number of exposures and subsequently, recognition of learned stimuli depends on the quality with which stimuli were encoded. However, to test this claim, I have tried to link performance during learning and testing in all the experiments described here. The time course of consolidation for declarative memories was tested using an fMRI study and the authors found that 30 days post learning, memories were cortical and considered to be stored in long-term memory (Takashima et al., 2006). With this in mind, I have tested the fourth claim by testing recognition memory at least 30 days post learning. In order to test the fifth claim, I have tested specific hypotheses regarding the neural correlates of memory using both electroencephalography (EEG) and functional Magnetic Resonance Imaging (fMRI) experimental paradigms.

Lastly, I was also interested in testing the eighth claim. As highlighted in the earlier section on STDP, I was enthusiastic about trying to apply computational models to explain mechanisms of memory. Since I love to do experiments and understand the results, and given how biologically relevant STDP models of learning are, I have tried to understand the results of my experiments in terms of predictions from the model. If observed behavior matches model predictions, it can be inferred that this is the actual mechanism by which information is processed and stored in the brain.

1.4 What do we know about the auditory system?

A) Ascending and descending auditory pathways

The auditory system performs an amazing feat in converting mechanical energy contained in sound waves into perception of our acoustic environment. The complexity of the neural mechanisms behind this ability becomes apparent when we start studying the steps in converting these mechanical signals into neural code. Additionally, preserving high temporal precision of the incoming sound is critical for localizing the source of a sound in space. The value of sound localization is apparent when we consider that all mammals, with the exception of mammalian subterranean species such as the pocket gopher and naked mole rat display sound localization capability to a certain degree (Heffner and Heffner, 1992). The auditory environment in subterranean habitats differs markedly, with restricted sound propagation. Other mammals can rapidly orient attention, gaze and behavior in response to a sound, implicating sound localization ability as a source of positive selection in evolution (Heffner, 1997). Sound localization involves the calculation of inter-aural differences, both in time of arrival and intensity of spectral components, both of which require high resolution temporal information to be preserved. This idea is further supported by the recent finding that sampling of acoustic stimuli is continuous and discretization of the signal happens at the perceptual level. This is in contradiction with findings in the visual domain where information in the environment is sampled using discrete snapshots. In fact, several mechanisms are in place for subsampling visual input at the source – (i) *microsaccades,*

or the microscopic movements of the eyes to "refresh" the retinal image of stationary stimuli and prevent adaptation, have been shown to correlate with perceptual alterations (van Dam and van Ee, 2006), (ii) *saccades*, or rapid eye movements to fixate different parts of the visual world, help to voluntarily and discretely subsample the visual environment, and (iii) cortical and thalamo-cortical *oscillations* which modulate attention have been proposed to contribute to discrete visual processing (VanRullen and Koch, 2003). On the other hand, it seems that auditory subsampling at the source renders sounds completely un-recognizable (VanRullen et al., 2014; Zoefel et al., 2015). These findings suggest that information in the auditory domain is preserved with high fidelity, at least until subsampling of acoustic information at the perceptual level.

The first step in converting mechanical sound signals into electrical based neural code happens in the cochlea, a structure in the inner ear. Before reaching the cochlea, sound waves are funneled into the external auditory meatus by the pinna to reach the ear drum. This mechanical signal is amplified and transmitted to the cochlea via the bones in the middle ear. The fluid filled inner ear contains the vestibular (semi-circular canals) and auditory (cochlea) sensory organs. After conversion into electrical signals (if the sound is above a given threshold of amplitude), sound information is carried into the central nervous system via the auditory nerve. The parts of the ear are shown in figure 1.12.

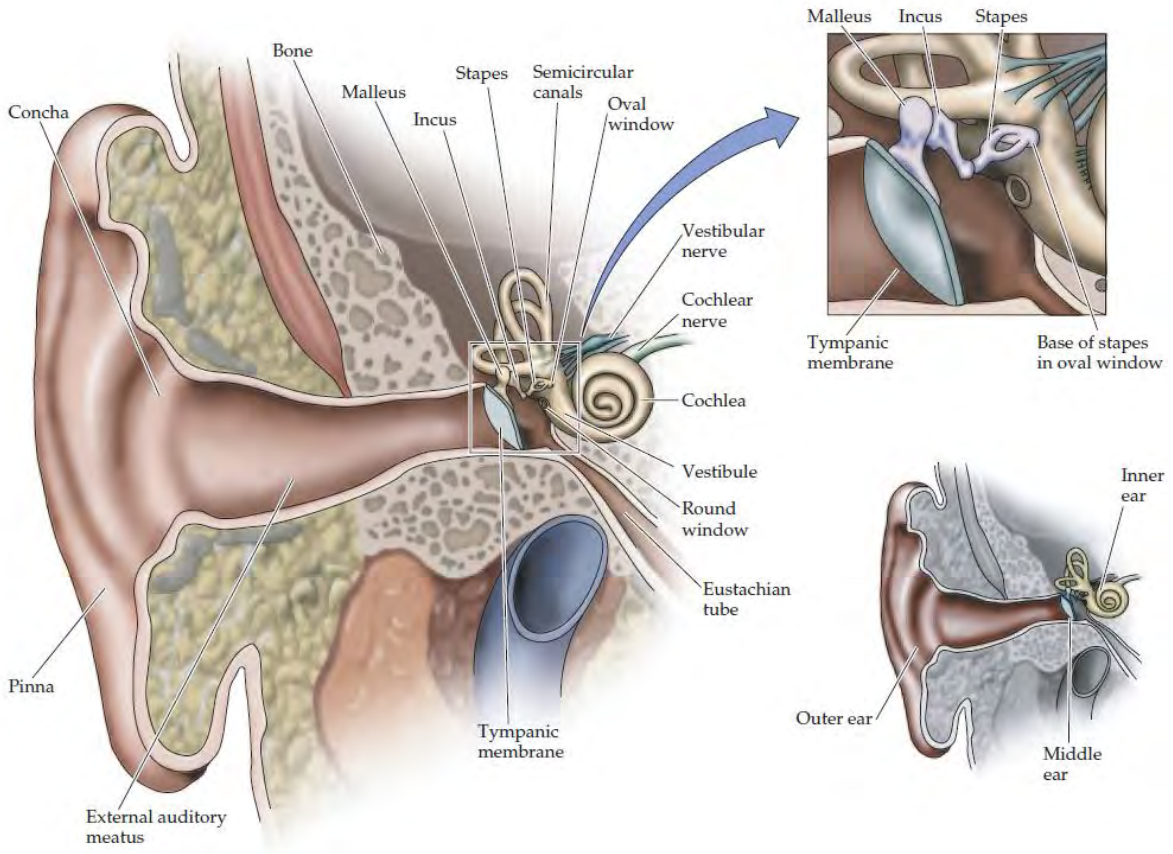


Figure 1.12: Main – illustration of the outer, middle and inner ear. The inner ear contains the cochlea, the auditory sensory organ converting mechanical sound waves into electrical spikes for processing by the central nervous system. Figure adapted from (Purves et al., 2008, chapter 12).

The cochlea has a complex morphology as illustrated in figure 1.13 [(Raphael and Altschuler, 2003) fig4 and fig 6 combined elegantly]. The two and a half turns of the cochlea contain inner hair cells that respond to sound waves. Movement of stereocilia, the apical modifications of inner hair cells, in response to incoming sound waves opens ion channels, effectively converting acoustic signals into action potentials (Engstroem et al., 1965). This results in all incoming sound waves to move from the base towards the apex of the basilar membrane. The basilar membrane is tuned to a range of frequencies

with the base responding optimally to high frequencies and the apex responding to low frequencies. The maximum amplitude of this 'traveling wave' corresponds to the frequency of the input sound, creating a tonotopic map of the incoming sound (Von Békésy, 1970).

Complex and fascinating as the study of the cochlea is, there are two aspects of cochlear function that are very interesting to note in the context of the series of experiments described in this thesis. The first is that ion channels in stereocilia are partially open when stereocilia are in their *resting* position. Therefore, there are spontaneous signals arising in the cochlea even in the absence of a sound, resulting in inherent noise in the system. Secondly, as is evident from figure 1.12, the cochlea receives efferent inputs from the lateral superior olive, which is a part of the superior olivary complex (discussed later in this section), in the brainstem. These efferent connections mediate the activity of the inner hair cells via the outer hair cells that contain several different neurotransmitters; figure 1.13 (Engstroem et al., 1965; Raphael and Altschuler, 2003). These two observations make it clear that even at the level of the cochlea, understanding how information is processed and features extracted is not simple.

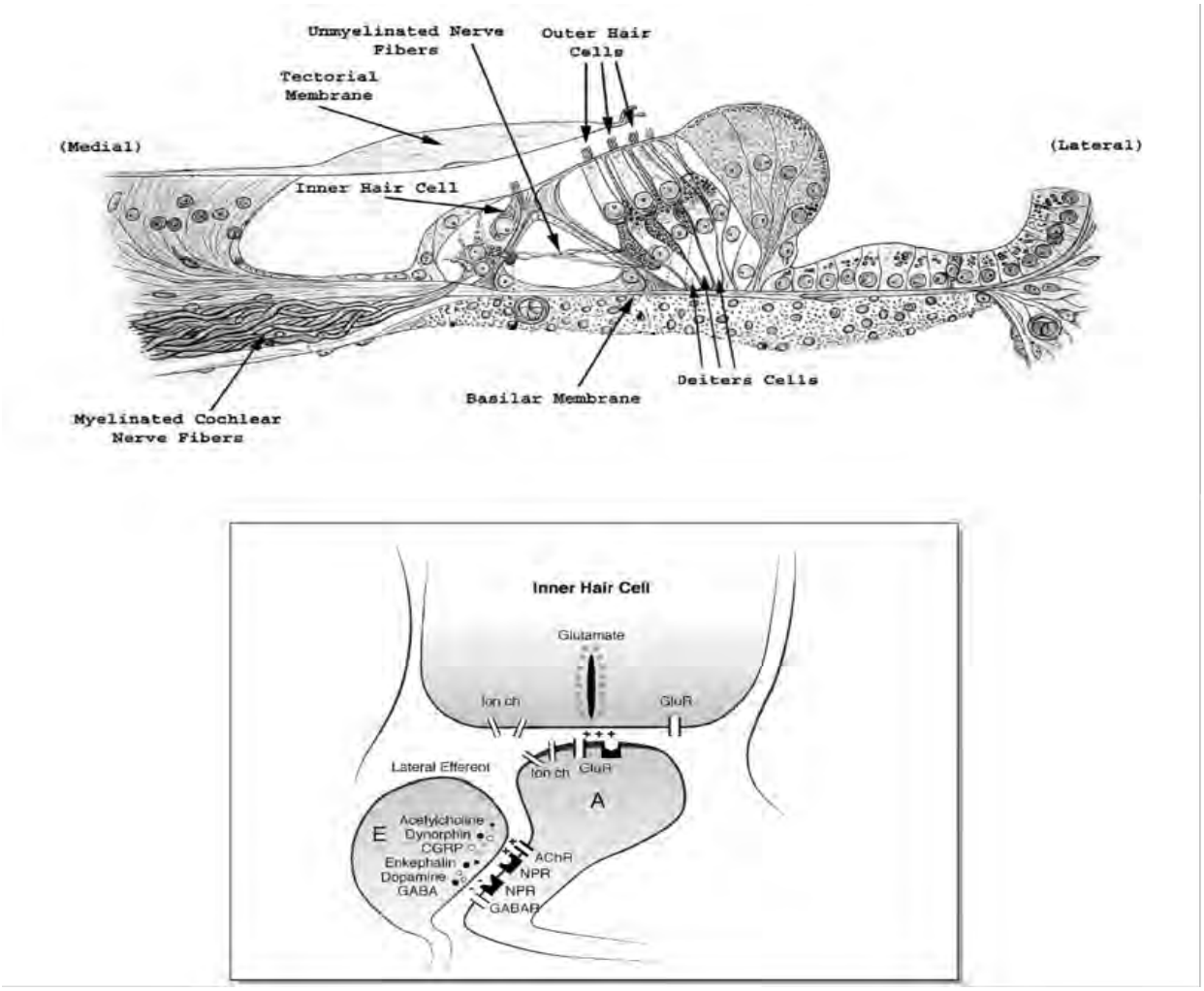


Figure 1.13: Top – illustration of cross section of a cochlea, highlighting the close physical proximity of the inner hair cell (sensory afferent) and modulatory outer hair cells (efferent). Bottom – illustration of this modulation via several different neurotransmitters at the synaptic level. (Raphael and Altschuler, 2003)

Next, the auditory nerve carries this sound information into the dorsal and ventral divisions of the cochlear nucleus and up to the auditory processing pipeline. After entering the brain, acoustic information travels along the ascending auditory pathway illustrated in figure 1.14.

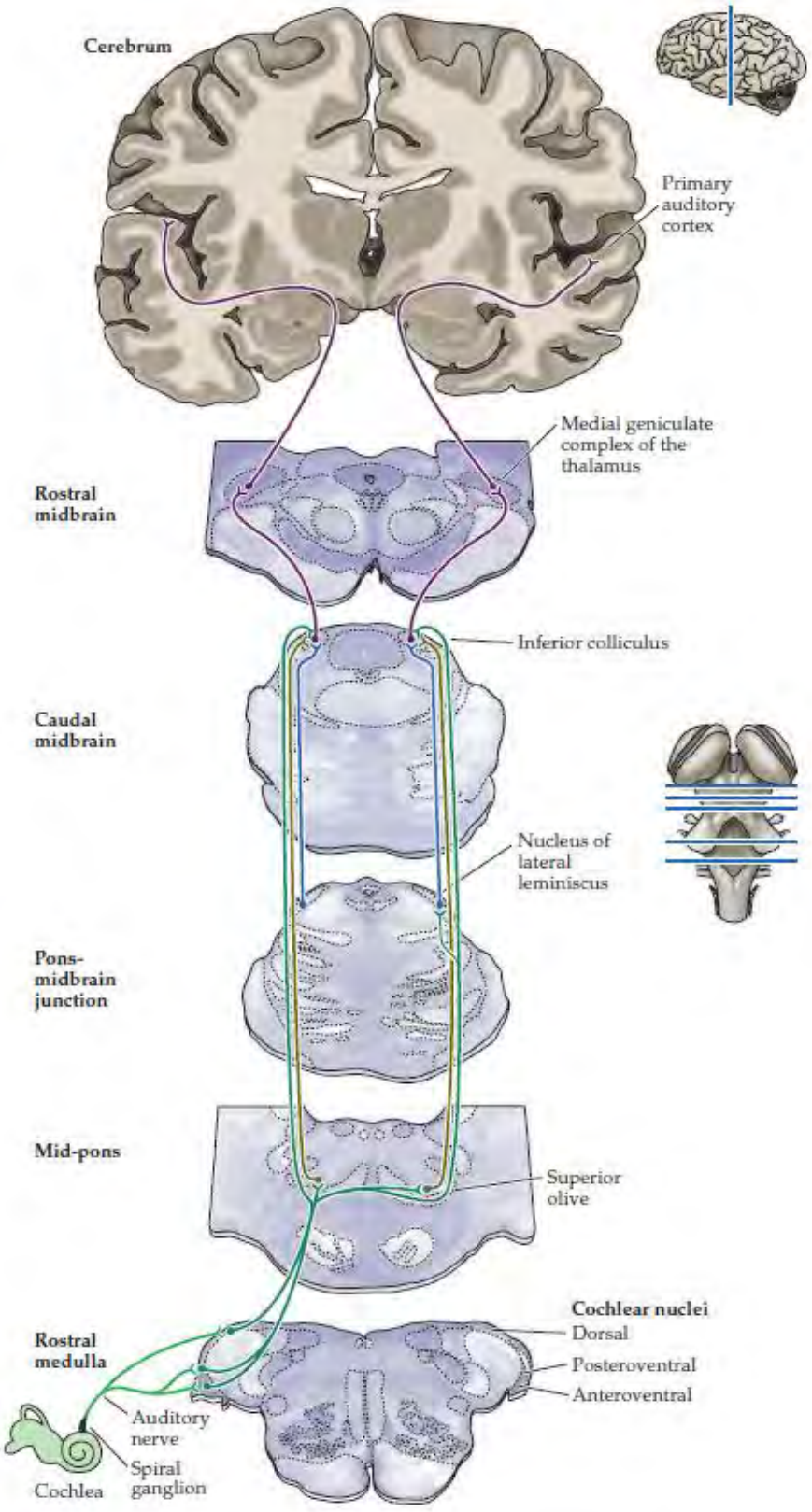


Figure 1.14: Ascending auditory pathway. Adapted from (Purves et al., 2008).

The cochlear nucleus complex, as shown in figure 1.15, is made of dorsal and ventral subdivisions. Based on the diversity of cell types and their computational properties, neurons in the dorsal cochlear nucleus have been hypothesized to extract multiple features from the incoming sound for parallel lines of processing. In fact, the dorsal cochlear nucleus has at least five different cell types – (1) pyramidal cells that are either type IV or giant cells which are the principal excitatory cells, (2) vertical or type II inhibitory inter-neurons, (3) granule cells which are excitatory cells receiving both auditory and non-auditory input, and (4) small cells (Olszewski and Baxter, 2014; Young and Oertel, 2003).

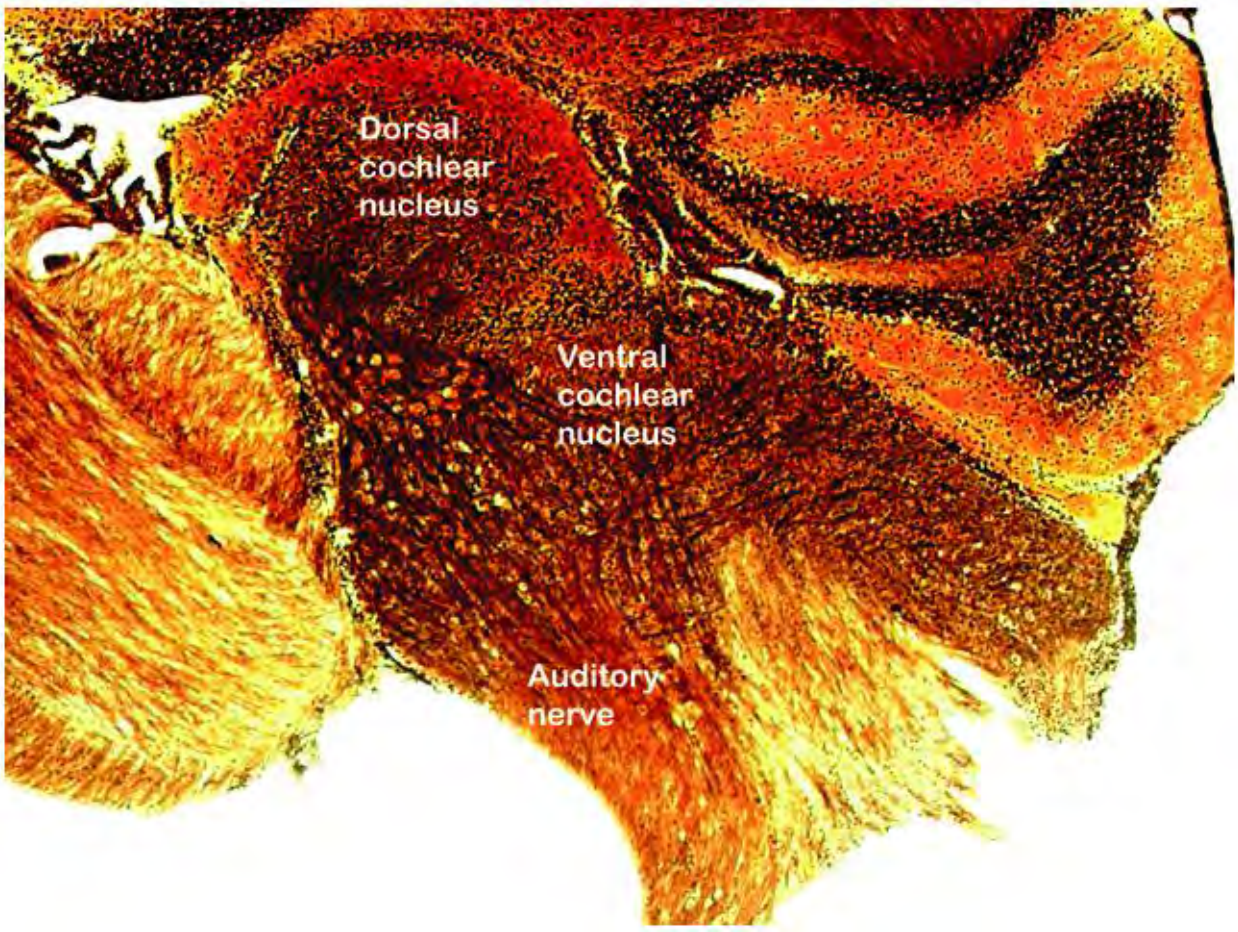


Figure 1.15: Histological section of the cochlear nucleus. Illustration adapted from the Professor Oertel'slab website: http://neuro.wisc.edu/faculty/oertel.asp.

These neurons are morphologically diverse and perform computations such as extraction of spectral components of complex sounds, a major step in sound processing, despite being in the periphery of the auditory system (Nelken and Young, 1996). Other neurons perform computations regarding periodicity and timing of features present in incoming sounds, by acting as coincidence detectors of input from the auditory nerve (Golding et al., 1995). Interestingly, an fMRI study (with cardiac gated acquisition to reduce physiological noise) on the role of the cochlear nucleus in humans found that neurons in this region code for temporal regularity/periodicity information (Griffiths et al., 2001). Therefore, a major function of the (dorsal) cochlear nucleus has been proposed to be periodicity detection and coding (Frisina, 2001).

On the other hand, the ventral cochlear nucleus contains predominantly two types of cells, the spherical and small neurons which project to the superior olivary complex, carrying precise temporal signals which are subsequently used to calculate inter-aural time differences. The ventral cochlear nucleus also contains multipolar and octopus cells near the junction with the dorsal cochlear nucleus and are believed to provide inhibitory input to some of the neurons in the latter.

The primary function of the superior olivary complex has been proposed to be sound localization (Olszewski and Baxter, 2014). In order to localize sounds in space, two variables are critical: time and intensity. The difference in arrival times of a sound in each ear gives an estimate of the direction of the sound source and the difference in intensity of the same sound in each ear gives an estimate of the distance of the sound. This is illustrated in figure 1.16.

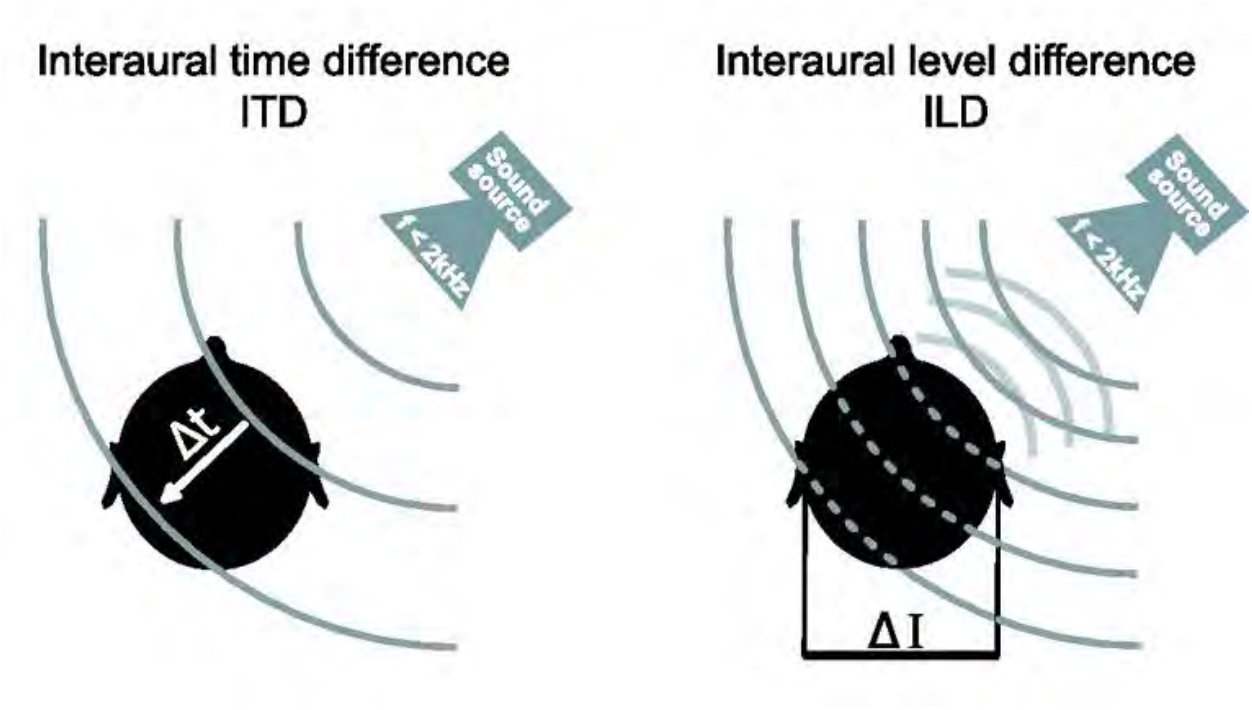


Figure 1.16: Schematic illustrating calculation of inter-aural time and intensity differences, primarily performed by the superior olivary complex. Figure adapted from (Grothe et al., 2010)

The superior olivary complex gets binaural inputs from the ventral cochlear nuclei. Computations of inter-aural time and intensity differences are achieved through two types of binaural input neurons – excitatory-excitatory neurons and excitatory-inhibitory neurons (Goldberg and Brown, 1969). Neurons in the medial subdivision principally calculate inter-aural time differences and neurons in the lateral subdivision principally calculate inter-aural intensity differences.

The next node in the ascending auditory pathway is the inferior colliculus, which forms a part of the tectum, a structure visible on the posterior surface of the brain. The inferior colliculus gets input directly from the dorsal cochlear nucleus as well as from the superior olivary complex before sending it onto the ipsilateral medial geniculate body

and the auditory cortex, effectively acting as a brainstem last relay. Three subdivisions of the inferior colliculus exist, each of which perform different computations. The *central* inferior colliculus gets binaural tonotopic input from the cochlear nuclei and the superior olivary complex. The *external* inferior colliculus is a multisensory region with visual, somatosensory, trigeminal and cortical afferents. This region is believed to be involved in the auditory-motor reflex pathway. Interestingly, the *dorsal* inferior colliculus gets contralateral monaural input and the neurons here project to the secondary auditory cortex, functioning to regulate auditory attention (Olszewski and Baxter, 2014). The cross section of an inferior colliculus is shown in figure 1.17.

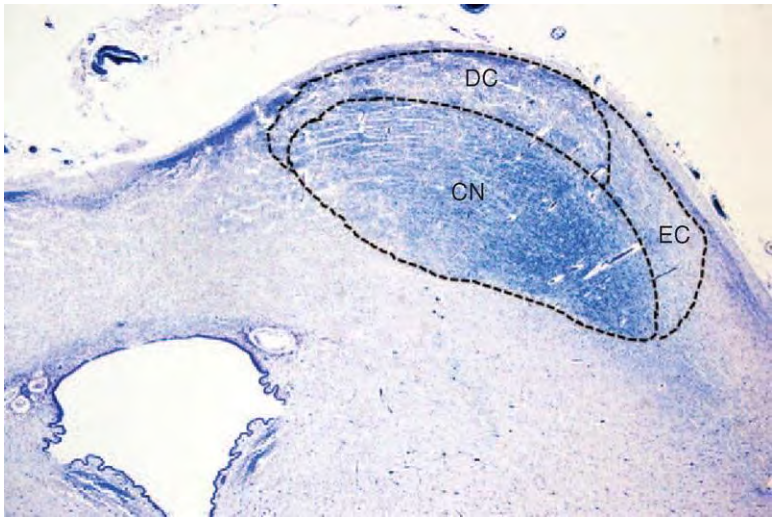


Figure 1.17: Histology and cross section showing the organization of the inferior colliculus in humans, using Nissl stain. Adapted from (Lavezzi et al., 2015).

Because of the diversity of morphology and function within the inferior colliculus, complex responses to stimuli have been observed with direct electrophysiological recordings. In one such study, authors used vocal stimuli rich in biologically relevant information such as identity and emotion. Authors found that some sub-populations

showed preferential activation to detect certain vocal features over others, while other sub-populations showed non-specific activations. Additionally, vocal stimuli elicited higher firing rates in the external inferior colliculus compared to the central and dorsal sub-divisions. In the same study, 78-100% of neurons responded to noise in all regions (Aitkin et al., 1994). Therefore, this nucleus plays an important role in processing information present in sound, and certain neurons in this nucleus seem to be able to selectively respond to important and relevant feature information present in sounds.

After exiting the brainstem, neural messages are processed by the auditory midbrain or the medial geniculate body (MGB) in the thalamus. Based on cyto-architecture, three major subdivisions have been observed in the human MGB, the *medial*, *ventral* and *dorsal* (Winer, 1984). The lateral geniculate nucleus, considered the visual analogue of the MGB, consists of magnocellular, parvocellular and koniocellular neurons that are clearly distinguishable in histological sections (Hickey and Guillery, 1979; Purves et al., 2008). The cellular morphologies of medial geniculate neuronal subtypes are less distinctive (Winer, 1984), as shown in figure 1.18. Functionally, however, the MGB has been shown to be critically important in sound processing. Unlike other auditory subcortical nuclei, strong *reciprocal* connections exist between the medial geniculate and the cortex (Kimura et al., 2003; Pontes et al., 1975). In order to understand how information is processed by this thalamo-cortical network, researchers simultaneously recorded cortical and medial geniculate neurons in the guinea pig. Using this paradigm, the authors were able to record from connected MGB - cortical neuron pairs that showed a very high correlation of activity, as well as from independent MGB and cortical neurons. Importantly, the MGB was seen to 'hold' (respond to) more

characteristic acoustic components present in sounds than corresponding (connected) cortical neurons. This was inferred to be a result of greater intra-cortical inhibition, resulting in greater selectivity of information processed by the cortex (Creutzfeldt et al., 1980). In other words, the internal representation of a complex stimulus was different in cortical and medial geniculate neurons - cortical neurons respond to a complex stimulus in terms of its simplest, representative form, while corresponding medial geniculate neurons were processing several component/characteristic features present in a complex stimulus. Further support for the idea that medial geniculate neurons respond to different features present in complex stimuli comes from the discovery of "combination sensitive" neurons in this nucleus (Olsen, 1994; Suga et al., 1998). In these studies, different sub-populations of the medial geniculate neurons were shown to have highly specific responses to species-specific vocalization sound signatures. These vocalization sound signatures usually contain complex spectro-temporal features and essentially, MGB neurons were able to differentiate different calls from each other.

These findings have been interpreted to be a result of non-specific and non-selective subcortical processing of acoustic stimuli, and researchers have hypothesized that real acoustic feature extraction occurs at the cortical level (Creutzfeldt et al., 1980). However, an alternate interpretation of the results discussed above is possible: that feature extraction depends on task and processing demands. When task demands require higher temporal precision of acoustic information – for example in tasks such as localization of or responding to sound sources – features extraction would necessarily occur at the subcortical level, where high resolution information is being held. As discussed in the sensory coding section, hierarchical processing of information leads to

successive layers storing features with increasing complexity (representative form) but lower resolution. From the studies discussed above, it is clear that medial geniculate neurons are processing multiple complex features present in incoming sounds and are able to differentiate component features with higher precision than corresponding cortical neurons.

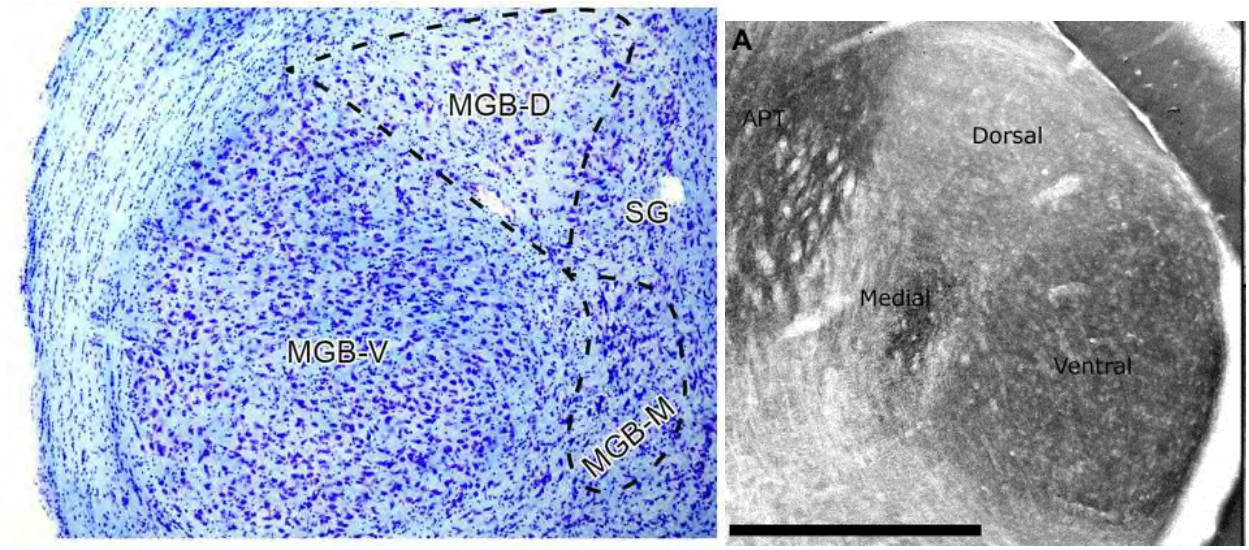


Figure 1.18: This figure illustrates the histological subdivisions of the medial geniculate body, visualized using different stains. Left: rat MGB visualized using Nissl stain, adapted from (Burianová et al., 2015). Right: mouse MGB visualized using cytochrome oxidase stain, adapted from(Anderson and Linden, 2011). The subdivisions are more apparent with cytochrome oxidase but under Nissl staining, MGB appears fairly homogenous.

After visual information enters the cortex, two processing pipelines are in place to decode "what" objects are in space and "where" they are with respect to egocentric co-ordinates. Information in the "what" or ventral visual pathway is accessible to conscious perception but not information in the "where" or dorsal visual pathway (Goodale and Milner, 1992). A meta-analysis of auditory imaging studies revealed the

existence of similar dual-pathway processing of auditory information, as shown in figure 1.19.

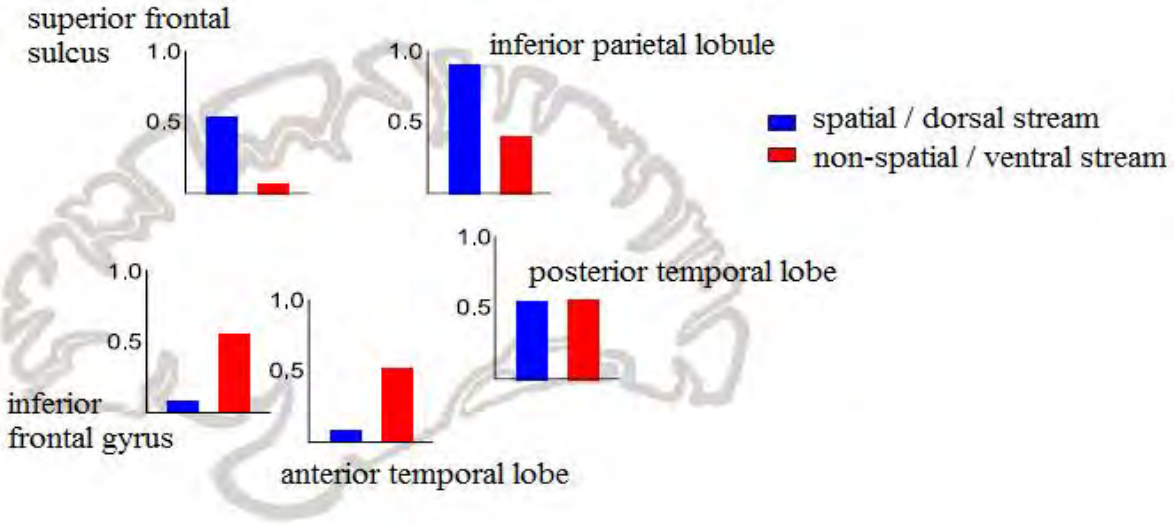


Figure 1.19: This figure shows the results of a meta-analysis of imaging studies implicating various cortical areas in auditory processing and the proportion of 36 studies (11 spatial and 27 non-spatial since some studies contained both spatial and non-spatial processing) demonstrating involvement in a particular area. This figure is adapted from (Arnott et al., 2004) and demonstrates the dual processing pathway in audition.

In the visual system, processing in the "what" auditory pathway is clearly accessible to consciousness but computations in the "where" pathway can proceed without any awareness, as suggested by work on patients with damage to the ventral processing pathway being able to respond to objects in their environment (scaling grip size to the orientation of the object) without any conscious perception (Goodale and Milner, 1992). A similar dissociation has been suggested in auditory processing pathways (Arnott et al., 2004). Further, cortical processing of sounds also seems to follow a hierarchical organization - in a study looking at the neural correlates of temporal pitch processing in the cortex, authors observed that with increasing complexity of the melody, activation

associated with processing of the sound was observed further away from the primary auditory cortex (Patterson et al., 2002). These studies show that cortical processing of sounds occurs in a hierarchical and task dependent manner.

So far in this section, the emphasis has been on describing the ascending auditory pathway. In general, subcortical ascending pathways are primarily sensory afferents and descending subcortical pathways are motor efferents. However, as described earlier, efferent fibers are present even in the cochlea figure 1.13 indicating that there is bi-directional flow of sensory information from and to the auditory periphery. The primary function of this descending auditory pathway, the *corticofugal* system, has been identified as shaping response properties of neurons in the ascending auditory pathway (Suga et al., 2000). Modulation of the incoming signal can be temporary, serving to amplify certain parts of incoming signal, also called adaptive gain control. Long term modulation of response properties of neurons imparts plasticity to these neurons. Descending modulation involving adaptive gain control and plasticity has been observed in both the auditory (Yan and Suga, 1998, 1999; Zhang et al., 1997; Zhang and Suga, 2000) and somatosensory systems (Ergenzinger et al., 1998; Krupa et al., 1999).

Cortical projections that descend from the auditory cortex have been shown to project to the ipsilateral MGB. Some of these projections are also directly connected with all bilateral auditory brainstem nuclei, including the inferior colliculus, the superior olivary complex and the cochlear nucleus. Notably, all the cortical descending projections are tonotopically organized and help modulate response properties of corresponding neurons in lower areas. Besides descending projections from the cortex,

some neurons in the inferior colliculus are also connected with olivo-cochlear neurons in the superior olivary complex that in turn project to the contralateral cochlea. These efferents form the outer hair cells seen in figure 1.13 (Suga et al., 2000).

A last point to note regarding the organization of the auditory system is the lateralization of function. As evident from the discussion above, several nuclei along the auditory pathway get binaural input from the ascending and descending projections. Despite this, pitch and speech processing was shown to be lateralized in a PET study (Zatorre et al., 1992). Further evidence for the lateralization of function in auditory processing comes from a mismatch negativity study of patients with unilateral lesions. When presented with tones monaurally, the mismatch negativity was only disturbed when stimuli were presented contralateral to the injury, implying that sensory memories are stored in the contralateral cortex. Patients with dorsolateral prefrontal cortex lesions had reduced mismatch negativity for stimuli presented in either ear indicating that this region influenced storage of sensory memory bilaterally (Alain et al., 1998).

From the above studies, it is evident that nuclei along the auditory pathway perform sophisticated computations to decode the spectral and temporal information present in complex sounds. The descending corticofugal system seems to modulate the perception of incoming stimuli. A question that arises is how and where auditory sensory memories are stored in the auditory pathway. These are discussed in the following section.

B) Sub-cortical plasticity in the auditory modality.

Several studies have implicated the auditory cortex in storing sensory memories (refer to section on mismatch negativity studies and models of sensory memory). Additionally, the N1 component, a large negativity seen about a 100 ms after sound onset in electrophysiological (EEG) studies, has been localized to sources in the primary auditory cortex (Reite et al., 1994). The amplitude of the N1 component was shown to change with learning acoustic features in a sound (Andrillon et al., 2015). Together, these studies show that primary auditory cortex may be involved in storing acoustic features. Given the importance of subcortical processing in audition and the dual nature of perception and memory, interesting questions regarding the role of subcortical structures in *storing* acoustic features come into focus.

The descending corticofugal system discussed in the previous section has been hypothesized to play a role in plasticity of subcortical auditory structures. In a seminal study, researchers electrically stimulated specific sites in the auditory cortex of the bat and observed the response of descending projections to the MGB and inferior colliculi. Stimulating a few neurons in the auditory cortex resulted in enhanced activity in MGB and collicular neurons coding for similar information while depressing neurons coding for other information (Yan and Suga, 1996). That is, cortical neurons with a preferential response to some frequency or spectral characteristic, when stimulated artificially, enhanced the activity of neurons in the MGB and IC with similar response preferences while inhibiting others. This study highlights *online* plasticity in the subcortical system, which probably helps to enhance perception and reduce internal noise.

The capacity of MGB neurons to store information was demonstrated when electrical stimulation of MGB neurons resulted in LTP effects in the MGB-Amygdala pathway in rats, effectively resulting in memory (Clugnet and LeDoux, 1990). Several other studies provide support and context for this finding. One such study showed that subgroups of MGB neurons are differentially activated and deactivated in response to the same complex sound features, suggesting that these features can be learned at the level of the MGB (Tanaka and Taniguchi, 1991). This regional selectivity for certain acoustic features over others has been hypothesized to correspond to the cortical area the neurons are reciprocally connected to (Buchwald et al., 1988). The most compelling evidence for memory in the MGB comes from the finding that single neurons in this nucleus can modify their response preferences following training to respond to a conditioning stimulus (Edeline and Weinberger, 1991). Furthermore, the descending projections from primary auditory cortex were demonstrated to directly modify response properties of MGB neurons (Villa et al., 1991). These results led Villa and colleagues to propose that the function of the corticofugal system is to filter the information access into cortical areas, or the "adaptive filter theory". By extension, therefore, these projections also help information storage at the cortical and subcortical levels. These studies demonstrate that computations of auditory perception are plastic in the MGB. Perhaps the combination sensitive neurons acting as coincidence detectors (Olsen, 1994; Suga et al., 1998) help to store complex acoustic feature ensembles.

Plasticity has also been demonstrated in neurons of the inferior colliculus. In one study, some neurons in the inferior colliculus were shown to differentiate complex sounds and their reversed versions, but others responded identically to both versions (Syka et al.,

1998). This demonstrated that some IC neurons code for features of sounds that are resistant to reversal. That is, some neurons might code for features of sound that are unchanged when a reversed version is played, and other neurons code for different features. Besides demonstrating memory in the inferior colliculus, this study showed that neurons in this nucleus code for multiple complex features.

It is unknown if the cochlear nucleus, particularly its dorsal subdivision (DCN), is capable of storing any kind of information over long periods. What is clear is that neurons in this nucleus are able to flexibly integrate spectral features present in a sound, as well as filter frequency components of incoming stimuli, resulting in the "dual nature of the DCN circuit". It is believed that these calculations are based on spectro-temporal receptive fields as previously defined (Aertsen and Johannesma, 1981). Therefore, while long term plasticity or memory might not be relevant to study in the cochlear nucleus, neurons in the dorsal subdivision demonstrate online plasticity (Young and Nelken, 1998) and this is interesting to explore further.

Finally, compelling evidence for plasticity in the subcortical auditory processing pathway comes from studies measuring the auditory brainstem response (ABR). ABR, was first discovered in 1970 and was thought to be useful to apply to clinical research (Jewett et al., 1970). The ABR is the differential activation between the mastoid and central electrode, seen within 10-20 ms after sound onset as shown in figure 1.20 (Skoe and Kraus, 2010). The characteristic amplitude and frequency components seen in the ABR are thought to reflect subcortical processing of these stimuli. While animal studies support the idea that subcortical learning does happen, this plasticity was assumed to

be short-term until (Russo et al., 2005) showed training-associated neural plasticity at the level of the brainstem in humans, specifically a post-training[7] improvement in detecting speech sounds within a noisy background correlated with changes in the ABR between 12 to 40 ms after sound onset. This indicates that post-training processing of a stimulus is different from pre-training in auditory brainstem nuclei.

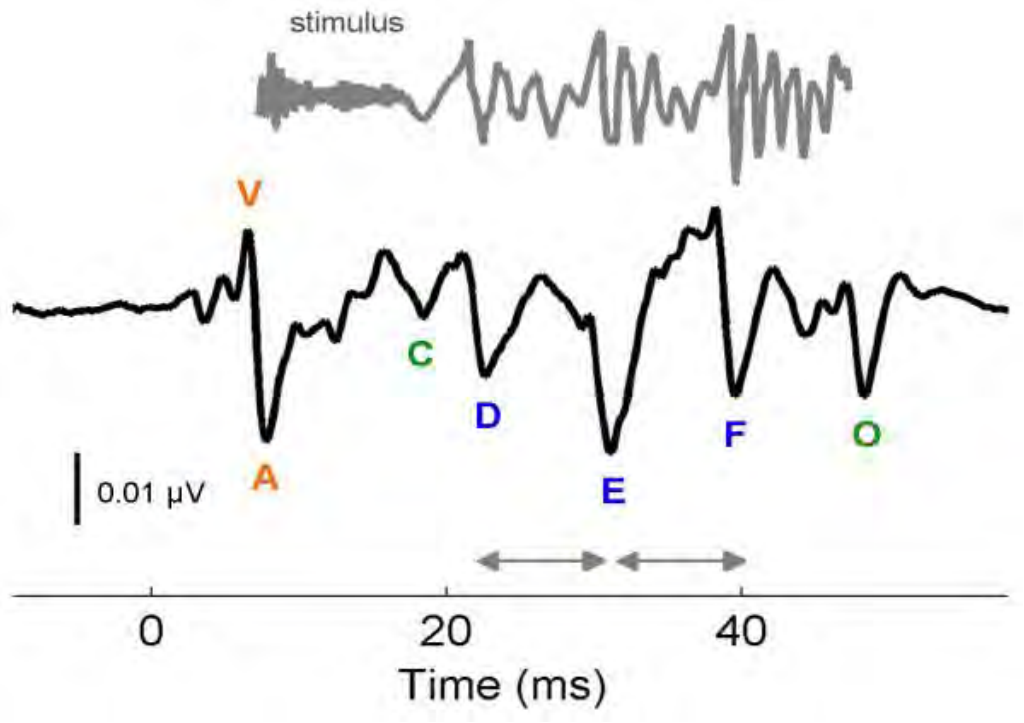


Figure 1.20: **Transient and sustained features in the ABR to a sound, /dɑ/ :**stimulus /da/ (gray) and response (black). The ABR includes both transient and sustained response features. This stimulus evokes seven characteristic response peaks that authors have termed V, A, C, D, E, F and O. These peaks relate to major acoustic landmarks in the stimulus. Peaks occur approximately 7 to 8 ms after the corresponding stimulus landmark, which is consistent with neural transmission time between the cochlea and rostral brainstem. V, A, C and O are considered transient responses in that they correspond to transient stimulus features. The V-A complex is often referred to as the onset response. The region between D and F forms the frequency-following response (FFR). Peaks D, E, F and the small voltage fluctuations between them correspond to the fundamental frequency (F_0) and harmonics of the incoming stimulus. Here, the stimulus plot is scaled to match the size of the response. Figure from (Skoe and Kraus, 2010).

[7] In this study, training consisted of a commercial audio training program that uses interactive games to improve language processing skills. (Earobics. http://www.earobics.com/. Evanston, IL: Cognitive Concepts.)

C) Role of auditory training on memory

It is clear from the previous section that some, if not all, subcortical nuclei display plasticity to a certain degree. Behaviorally, training results in a gradual improvement in performance until a plateau is reached. Extending these findings, it can be hypothesized that certain kinds of auditory training, such as training to discriminate pitch, would also modulate the response of these neurons until learning/expertise is achieved. Factors such as musical expertise are the result of training and are known to influence auditory perception and memory. Musical training is accompanied by neural changes including structural, functional and connectivity changes (Strait and Kraus, 2014). Non-specific changes and improvements are also possible in the brainstem in response to training: prior training in music was shown to improve verbal (still acoustic based) but not visual memory (Ho et al., 2003). This shows that previous experience affects what features are stored, at least within a given modality. Auditory experience, and not only musical or verbal training in a particular domain, influences processing of information as evidenced by the continuous evolution of the auditory cortex (Pantev and Herholz, 2011) and brainstem (Song et al., 2008) through life. Based on these findings, it has been proposed that speech and music processing share common underlying mechanisms, and training in one would then cause performance in the other to improve (Besson et al., 2011).

1.5 What do we know about auditory sensory memory?

A) Using meaningless auditory stimuli to investigate memory

The advantages of using meaningless stimuli to investigate memory were highlighted by Ebbinghaus (Ebbinghaus et al., 1913). Further, as discussed in the preceding sections, understanding sensory coding mechanisms responsible for perceiving and encoding meaningless sensory stimuli will further our understanding of learning in infancy, and how sensory information in its most simple form is processed. Therefore, the discovery by Guttman and Julesz that participants can detect repeating features in auditory noise when cycled continuously back to back to create *cyclic noises (CNs)* was fortuitous. It demonstrates that people can have limited conscious access to features in meaningless information (Guttman, N., and Julesz, 1963). Another interesting aspect that emerges in light of these findings is that these Gaussian sounds seem to lie at the border between implicit and explicit processing. Participants are able to consciously detect that there are acoustic features that repeat, but are only able to experience this in terms of vague perceptual events such as "clanks" or "whooshes". That is, participants demonstrate clear behavioral biases in terms of recognizing cyclicity, but are somehow unable to access individual features directly.

Since then, numerous studies have investigated the properties of this capacity, as discussed below.

First, investigations of memory for meaningless sounds in animals reveal that the ability to store information based purely on the statistics of occurrence is fundamental and shared across species. Rodents when exposed to cycle lengths of 20,40,60,80 or 100 ms for 8 seconds were able to detect cyclicity, and best performance was observed for 20 ms cycle lengths (Kaernbach and Schulze, 2002). Using a 'go-nogo' paradigm, cats were also shown to be able to discriminate cyclic sounds, and showed a sharp decline in ability to detect repeating features when cycle lengths were greater than 450-500 ms (Frey et al., 2003). These studies highlight that auditory storage of meaningless information, just like sound localization capacity (discussed in section 1.4A), is not restricted to humans.

Behavioral responses to cyclic noises have been investigated using tapping studies. In these studies, participants were asked to tap along with the perceived rhythm of cyclicity. Using this paradigm, it was shown that tapping in response to a given feature is consistent within participants. That is, a participant consistently recognized and responded to the same feature during multiple presentations of the CN, suggesting that a particular segment of the whole is learned and detected reliably (Limbert and Patterson, 1982). The authors further investigated these results by analyzing the spectrogram of the sound with respect to tapping behavior, and did not find any correlation between peaks in the short term spectrogram and the tapping point. This suggests that no distinguishing features on the amplitude spectrum of Gaussian sounds are preferentially detected or learned by participants. In another study, tapping performance for a given CN was compared between participants (Kaernbach, 1992). Here, the author noted that different participants notice features at different times – or

different temporal features - even if the consistency of tapping remained within participants and the performance was equivalent across participants.

Another set of studies have tried to investigate this phenomenon in order to understand the limits and capacities of auditory perception and working memory. One of the earliest studies with this objective used gaps in patterns of cyclic noise segments. In this study, the author asked participants to listen to two noise patterns separated by a gap in each trial and judge if the first and second patterns were identical. The author varied the gap duration (between 0.5 to 8000 ms) and pattern length (between 2 to 500 ms). Pollack demonstrated that a gap duration of 0.5 s and a pattern duration of 64 ms were optimal for detection of CNs (Pollack, 1972). These results were interpreted to imply that a salient gap of half a second serves to enhance features present in successive patterns. The author also observed huge inter-individual variations and could offer no explanation for it. In another study, the tapping task was used with cyclic noises of different lengths, and the maximum size of individual spectro-temporal features that were perceived and stored in Gaussian noise was found to be 100 ms (Kaernbach, 1993). Interestingly, this was the first study to define acoustic features in noise as complex spectrotemporal segments involving multiple and variable auditory channels for processing. A third study explored the role of training on the capacity to detect features in Gaussian noise. When the length of the cyclic segment exceeds two seconds, participants start declining in their ability to detect cyclicity. However, by training participants with longer lengths of the cyclic segment, Warren and colleagues demonstrated that with strong encoding, participants were able to detect cyclicity in 10, 15 and in some cases even 20 second long cyclic segment. To ensure strong encoding,

cyclic noises were presented with a synchronous visual light cue and participants were instructed to tap to the rhythm of cyclicity. After the light cue was turned off, however, tapping remained consistent even when the cyclic noises were made aperiodic using variable noise fillers. To make a cyclic noise aperiodic, instead of presenting the "frozen" segment of noise back to back, the authors introduced a variable noise segment between successive presentations of the frozen segment (Warren et al., 2001). This study hints at two characteristics of memory for Gaussian noise segments: segments once learned are detected in a stable and predictable manner and quality of encoding correlates with performance measures. Another study shed light on the robustness for the storage of individual acoustic features. Using a tapping paradigm adapted to participants' performance, Kaernbach investigated different aspects regarding the perception of noise. Participants were asked to tap along (once per period) to the perceived periodicity of the CN. Once participants had 8 consecutive taps that were correct, the trial was considered to be a success. The author was able to demonstrate that participants are unable to tap along to cycle lengths greater than 6 seconds (Kaernbach, 2004), without the aid of an external cue as used by (Warren et al., 2001). In another experiment, Kaernbach presented participants with a cyclic noise for 7 seconds and then an 8.25 second period of silence, followed by one presentation of the cyclic noise at the end of the trial. Participants had to indicate if the cyclic noise presented at the end was the same as the 7 second noise heard at the beginning of the trial. Participants were able to do the task, thereby demonstrating poor susceptibility to interference from highly similar patterns, suggesting that the resolution of features held in working memory and short term memory is very high and specific (Kaernbach,

2004). This finding is in line with predictions from an STDP model of perception and encoding as discussed in section 1.3B.

More recently, using an elegant experimental paradigm, Agus, Thorpe and Pressnitzer were the first to investigate long-term sensory memory using cyclic noises (Agus et al., 2010). In a cyclic, non-cyclic discrimination task, participants were implicitly presented with some exemplar cyclic noises, or *target CNs,* several times within a block. Over several experimental blocks, participants were thus presented with target CNs, novel CNs (heard only once throughout the experiment) and non-cyclic noise segments (Ns). The authors found that participants rapidly improved in detecting cyclicity in some, but not all target CNs. Those target CNs participants successfully learned to reliably discriminate were considered "learned" (figure 1.21, left). In addition to this, authors also investigated long-term implicit recognition for these sounds 2-3 weeks post-learning. When presented with target CNs that had been learned during the first session, participants were able to detect most of these sounds as cyclic from the first presentation during the second session. In comparison, detection of novel CNs was around chance (figure 1.21, middle). This is the first direct evidence that participants store meaningless information for longer than working memory stores allow. Lastly, to test the robustness of memory, the authors also presented participants with intact and reversed versions of the learned target CNs. Their results were very surprising; participants showed implicit recognition memory for intact and reversed versions of the target CNs (figure 1.21, right).

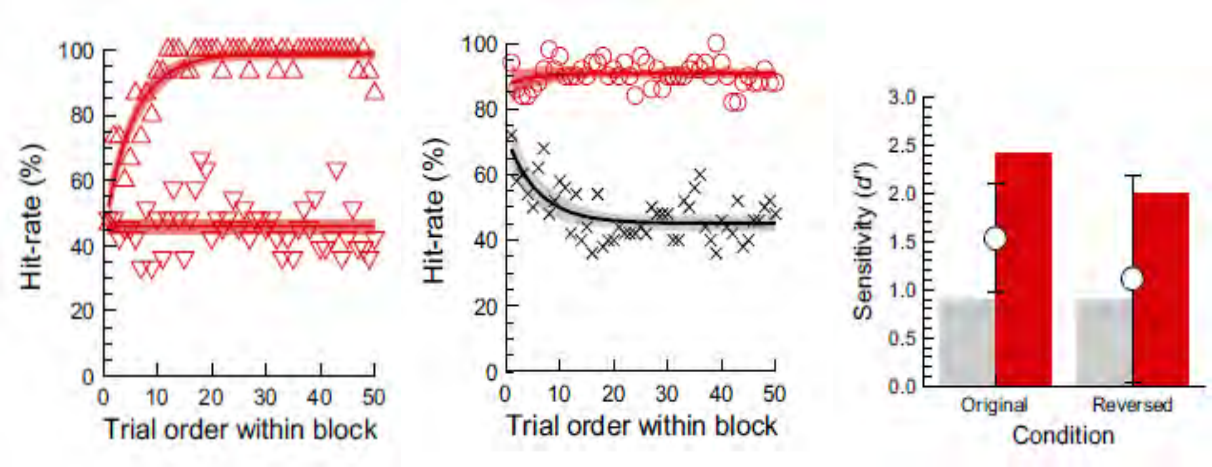


Figure 1.21: Summary of results from the first study investigating long term memory using cyclic noises (Agus et al., 2010). Left: discrimination performance during the learning sessions showed that some target CNs are rapidly learned, while detection of cyclicity for other sounds remains around chance across number of exposures. Middle: discrimination performance during the testing session demonstrated implicit recognition of learned target CNs (red) compared to novel CNs (black). This memory was significant from the first presentation during testing and was not a result of within-session improvement. Right: significantly better discrimination performance for intact and reversed versions of learned target CNs (red) compared to novel CNs (gray) during the testing session, indicate that memory for acoustic features in noise is robust.

In a follow up study using a similar implicit learning and recognition memory paradigm, in addition to target CNs and novel CNs, participants were presented with *mixed* stimuli during the testing session (Agus and Pressnitzer, 2013). These mixed stimuli were not cyclic: half the stimulus was made of novel noise segment that participants had never heard before and the other half of the stimulus was the learned segment of a target CN presented once. These mixed stimuli were by and large mistaken as cyclic, even when made aware of the existence of these types of trials, as shown in figure 1.22. These results link the perception of cyclicity to memory for features that are learned in target CNs, and demonstrate that once acoustic features are

encoded, they are recognized outside their encoding context. In another study investigating long-term memory for Gaussian noise, researchers introduced a temporal jitter during encoding. As a consequence, participants showed reduced sensitivity to learned acoustic features, further supporting the notion that quality of encoding influences the recognition memory for Gaussian noise (Rajendran et al., 2016).

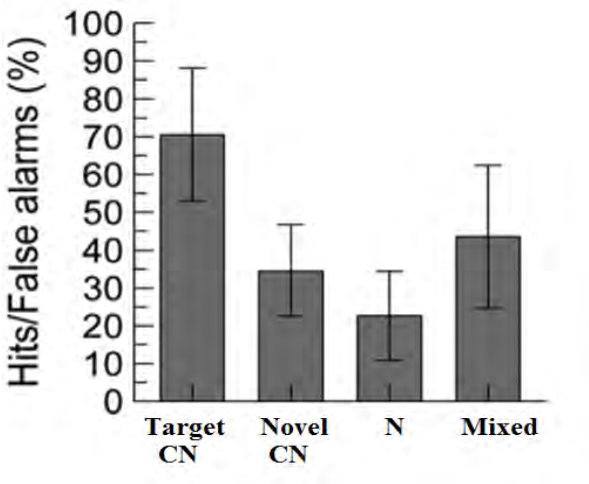


Figure 1.22: Results from the follow up study investigating long term memory using cyclic noises (Agus and Pressnitzer, 2013). Mixed stimuli were not cyclic but were mostly mistakenly classified as such, linking perception of cyclicity to features learned in noise.

The studies described above highlight certain properties of memory for Gaussian sounds – that acoustic features are robustly stored and the quality of encoding influences long-term recognition memory. In addition to these behavioral experiments, a few exciting studies have also investigated the neural mechanisms of memory for Gaussian sounds using electroencephalography (EEG) and functional Magnetic Resonance Imaging (fMRI) techniques.

Luo and colleagues investigated the role of the cortical sources in learning Gaussian sounds using EEG (Luo et al., 2013). By collecting surface brain activity while participants were discriminating target CNs, novel CNs and Ns, they demonstrated that phase of EEG oscillations arising from the auditory cortex were different for target CNs and novel CNs, suggesting that the auditory cortex is involved in storing learned acoustic features (figure 1.23, left).

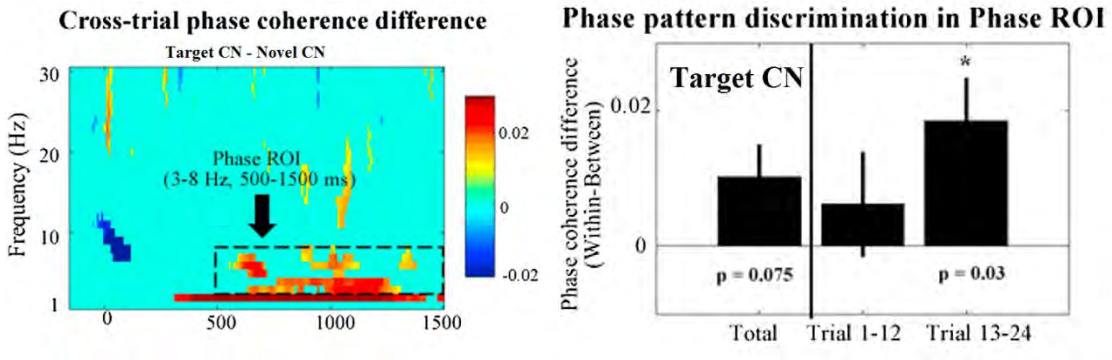


Figure 1.23: Results from an EEG study investigating the relationship between phase of oscillations and learning of target CNs (Luo et al., 2013). Left: observed phase coherence in the low frequency (delta and theta range) was reliably different for target and novel CNs about 500 ms after sound onset. Right: Individual target CNs were tracked differently in the second half of learning suggesting that after a certain number of presentations, activity in the auditory cortex was different for each target CN that is being learned.

Additionally, these authors also showed that the observed phase coherence for individual target CNs was different and evolved with learning (figure 1.23, right). That is, while learning, the phase coherence increasingly tracks a target CN with successive presentations and that this tracking is different for each target CN that is learned. Since EEG tracks cortical activity, these results demonstrate that individual features of

learned target CNs may be stored separately in the auditory cortex and observed phase coherence could reliably distinguish individual target CNs post-learning.

The rapidity of encoding features in cyclic noise was further explored using another EEG study (Andrillon et al., 2015). In this remarkable study, participants were presented with 200- or 500-ms patterns that repeated every 500 ms, embedded randomly in 8 minutes of continuous noise. Participants were instructed to detect changes in amplitude modulations, a secondary task that had nothing to do with paying attention to the CNs. A CN in this study was a segment of Gaussian noise that was cycled 5 times either continuously (500 ms CN presented 5 times) or sparsely (200 ms CN + 300 ms of novel N presented 5 times). While some of the patterns (novel CNs) were only heard once, others (target CNs) were embedded multiple times within the 8 minutes of continuous noise (figure 1.24, top panel, illustrates the stimuli). Amazingly, the authors demonstrated that fully developed evoked potentials were observed *rapidly*, within 5 presentations of a repeating pattern! The observation that ERPs develop within a few presentations, without any of the participants actually attempting to do so, suggests that even in the absence of a task, the brain can *automatically* learn patterns within a few exposures (figure 1.24, middle panel). Further, the amplitude of evoked potential increased with each successive presentation, providing evidence of 'online' learning mechanisms in the brain. That is, the change in neural activity as it learns patterns in a cyclic noise was, incredibly, quantified in terms of changes in the N1 amplitude. Overall, amplitudes of evoked potentials arising from central electrodes [believed to reflect activity in the primary auditory cortex (Reite et al., 1994)] reliably differentiated target CNs from Novel CNs and Ns. A correlation was found between

number of presentations and amplitude of evoked potentials (figure 1.24, bottom panel). These evoked potentials (in terms of amplitude, coherence and spectral power) also correlated with sound amplitude discrimination performance. Thus, using a simple experimental paradigm, the authors were able to demonstrate that patterns in CNs are encoded rapidly and automatically, as a function of number of presentations and without any top down attention.

Lastly, in an fMRI study using meaningless tone clouds[8] that were either cyclic or non-cyclic, specific hypotheses regarding cortical sources of memory for Gaussian sounds were tested (Kumar et al., 2014). Using multi voxel pattern analysis of activity in response to *individual target cyclic tone clouds*, the authors demonstrated that patterns of activity in the hippocampus and the auditory cortex (planum temporale) reliably distinguished individual target cyclic tone clouds. This study sheds light on the role of specific regions, particularly the hippocampus and the auditory cortical areas, in storing fine acoustic features.

Overall, these behavioral and neuroimaging results have raised several more questions regarding how and where acoustic features are perceived and stored in the brain. We have attempted to address some of these in the following series of experiments.

[8] . These stimuli were used instead of Gaussian sounds to ensure that the stimuli were distinguishable from each other within the context of the MRI scanner's gradient noise.

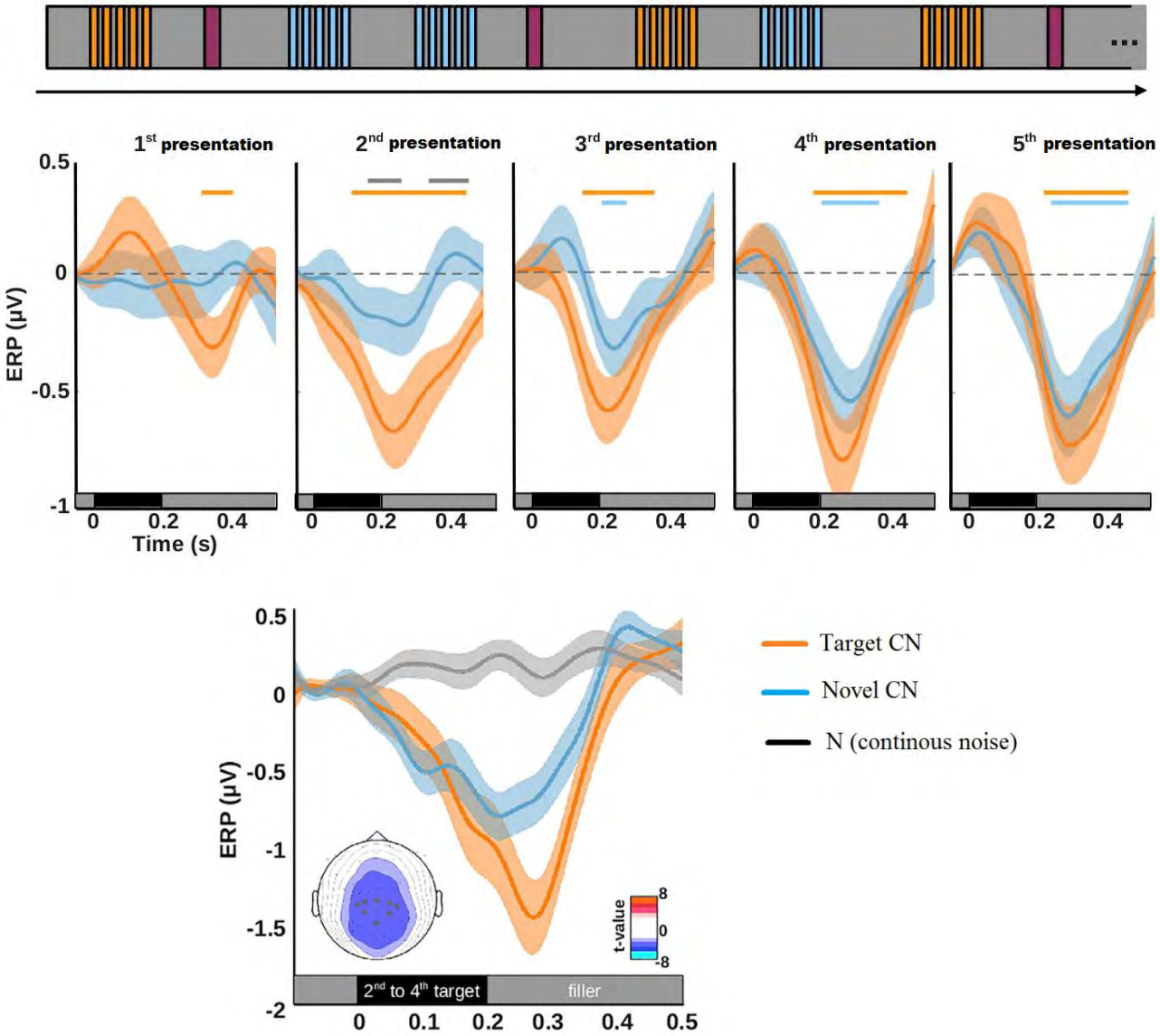


Figure 1.24: Results from an EEG study investigating the rapidity of encoding target CNs (Andrillon et al., 2015). Top: Sequence of a trial where 8 minutes of continuous noise included multiple presentations of target CNs (orange) and single presentations of novel CNs (blue). Participants were instructed to respond to amplitude deviants (red) Middle: Within a CN, reliable evoked potentials were observed within 5 presentations. Bottom: Overall, amplitude of evoked potential in central electrodes increased as a function of number of exposures to the cyclic segment.

B) Principal objectives and chronology of this thesis

As discussed in the previous section on using meaningless stimuli to investigate sensory memory, the findings of Agus and colleagues (Agus et al., 2010), namely that participants are able to store some information about these sounds over several weeks [figure 1.20], raise several fascinating questions. Most of these questions fall under one (or more) of 3 categories –

1) What are the characteristics of long-term memory for meaningless stimuli?
2) What are the mechanisms and neural correlates of this implicit recognition memory?
3) What neural codes are responsible for this learning and how well do behavioral data match existing models of learning and memory?

The existence of long-term memory for some but not all meaningless stimuli is in itself surprising, considering the fact that by definition, any storage of information about these stimuli would be purely a function of number of repetitions. And yet, as shown by the study highlighting poor memory for details on pennies (described in section 1.2C), lots of different kinds of information are not stored, despite a high probability of occurrence in the environment. We can therefore infer that at some point, a decision is being made about what information is consolidated and what information isn't. While we don't yet understand this decision process, a good way to start is by studying the characteristics of the information that is stored. One parameter that has been shown to affect such learning is attention, but we do not yet fully understand how the decision to delegate the limited attentional resources is made. In other words, in real world

scenarios, with stimuli competing for attention, how are relevant and important features 'chosen' to be processed and stored while others are forgotten and pruned? Beyond attention, several questions remained unexplored: what is the resolution of the information stored? How does prior expertise/training in the same sensory modality (e.g., musical training) affect recognition performance? How does sleep affect selective consolidation? By answering some of these questions, we can understand what features of the stimulus are stored and by inference why these are important, perhaps in terms of evolutionary drives. In my thesis, I have tried to address hypotheses that fall under all three categories using numerous psychoacoustic, behavioral and neuroimaging techniques.

I started my PhD in September 2013 and since then, I have tried to investigate as many of these fascinating questions as I could. Soon after I joined, in November, I started my first experiment by running a behavioral study, similar to that done by Agus and colleagues (Agus et al., 2010), with an aim of finding the smallest possible auditory feature that participants can learn. I picked 20- and 10-ms bin sizes to scramble learned sounds, fairly convinced that participants would treat these sounds as new. To my utter surprise, participants actually seemed to implicitly recognize such highly degraded versions of sounds, making me losing a bet with one of my thesis supervisors, Dr. Simon Thorpe, who (rightly) predicted that participants would show some memory for even highly scrambled sounds. This experiment and results are discussed in chapter 2. Inspired by such unexpected and interesting results, I set out to understand the spatio-

temporal correlates of this memory, using a similar experimental paradigm in a combined EEG-fMRI setting. Here I faced not insubstantial technical issues, and I spent about a year and a half of my thesis trying to run this experiment successfully and make sense of the data. As a novice to both EEG and fMRI, I was on a steep learning curve, and spent several months trying to choose the right parameters, settings, functions and toolboxes to analyze the data. Fortunately, I had help from Dr. Florence Remy, the co-director of my thesis and expert in all things fMRI, to help me out when I got really stuck. It was all worth it in the end, when I was finally able to analyze the data and again, to my surprise, subcortical areas, especially the medial geniculate body seemed to be heavily implicated in memory for meaningless acoustic features. These results are discussed in chapter 2. By then, I was already at the end of my second year with only one year left to explore all the follow up questions in my mind (and there were many), as well as writing up my thesis. In November 2015, we decided to implement and run a novel experimental paradigm testing several hypotheses at once. We were particularly interested in testing our hypotheses regarding the storage of very short acoustic features. The results are discussed in chapter 4. Armed with these fascinating results, I have spent the last few months of my thesis trying to understand the implications of our findings and what they have taught me about the mechanisms of sensory memory. I discuss these, along with all the interesting and wonderful questions regarding memory mechanisms that remain to be explored with meaningless/Gaussian sounds, in chapter 5.

II. Robustness of memory for implicitly learned Gaussian sounds

2.1 Introduction

Given the interest in understanding the characteristics of memory for Gaussian sounds, a first question we wanted to study concerned the temporal and/or spectral features in noise which are actually stored. Interestingly, when learned cyclic noises were played backwards during a retention test (Agus et al., 2010), detection of cyclicity was more accurate for reversed versions of learned sounds than for novel cyclic sounds, suggesting that acoustic features which are (implicitly) encoded are preserved in the reversed version of the sound. In this first experiment, we investigated how implicit memory performance varied when learned stimuli were modified using different acoustic transformations.

We also investigated the link between strength of meaningless stimuli encoding and subsequent memory performance. Turk-Brown at al. reported that brain regions such as the PPA (Parahippocampal place area, involved in scene memory) show higher activity during the encoding (first exposure) of scenes that were subsequently recalled vs. scenes that were forgotten (Turk-Browne et al., 2006) in both implicit and explicit paradigms. Performance in implicit encoding of meaningless sounds typically shows high inter-individual variability (Agus et al., 2010), and we were interested in exploring the relationship between strength of encoding and subsequent implicit recognition of these stimuli.

Putting these findings together, we hypothesized the following:

1) Long term, implicit memory for auditory noise (demonstrated as a preferential bias to detect cyclic features in learned vs. novel sounds) would be resistant to acoustic transformation, declining with increasing degree of transformation from the original learned noise.
2) This resistance to transformation would depend on how strongly the stimuli were encoded.

These claims were tested using an implicit encoding and subsequent long-term implicit recognition paradigm, as previously described (Agus et al., 2010). To test the first hypothesis, in addition to the old (learned) and novel sounds used in traditional memory retention tests, participants were presented with modified versions of the learned sounds. In some trials, the temporal origin of a learned sound was randomly shifted (looped CNs), changing the temporal expectancy of the learned feature(s) but preserving acoustic properties and surrounding context. In other trials, learned sounds were randomly shuffled to disrupt both temporal expectancy and surrounding context of learned features (scrambled CNs). To test the second hypothesis, implicit recognition performance on modified versions of learned sounds was considered in relation to learning performance. We predicted that implicit recognition during the retention test would vary as a function of acoustic transformation from the original learned sound. We also predicted that implicit long-term recognition would be higher in participants showing better performance during the learning session. These hypotheses regarding implicit recognition performance are summarized in figure 2.1. Our first hypothesis was

tested by comparing discrimination performance for intact vs. modified versions of the sounds as well as discrimination performance for intact vs. novel sounds. We predicted that novel and highly degraded sounds would elicit equivalent detection performance, around chance.

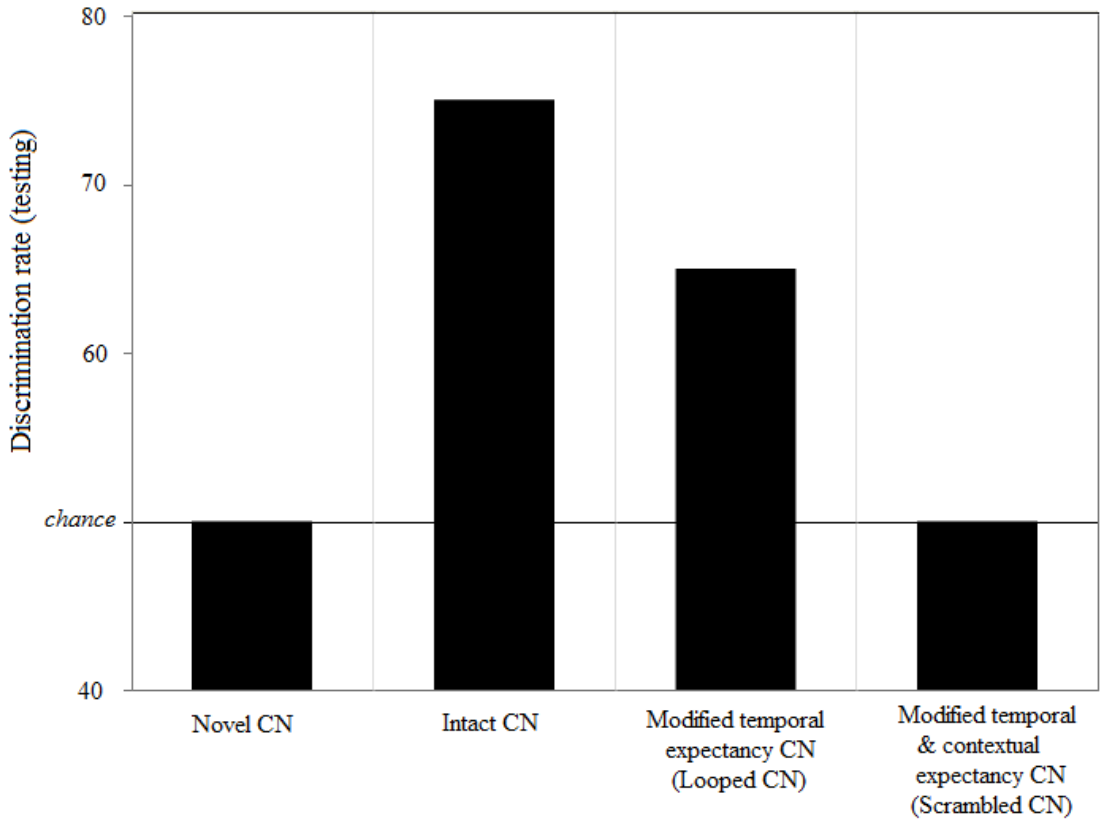


Figure 2.1: Predicted discrimination performance during the testing session: We predict that novel cyclic sounds and scrambled versions of the learned CNs would have equivalent performance, around chance. We also predicted that intact versions of the learned sounds would elicit higher discrimination performance than novel CNs.

2.2 Materials and methods

These hypotheses were investigated in one experiment with 2 versions. Participants were randomly assigned to one of the two versions of the experiment, which only differed in small aspects, as described in the procedure.

2.2.1 Participants

A total of 37 participants between 20 and 30 years of age, with self-reported normal hearing, were screened for the experiment. Twenty-five of these participants (mean age = 24.32 years, S.D = 3.07) were finally included. All participants were compensated for their time with gift cards pre-loaded with monetary values proportional to the extent of their participation, ranging from 10 euros (only screening) to 40 euros (completing both sessions of the experiment). They were instructed that the purpose of the experiment was to assess auditory discrimination and were naïve to the actual hypotheses of the experiment. All participants gave written informed consent in accordance to the declaration of Helsinki and the University of Toulouse and CNRS requirements for research with human participants [Protocol: CPP14-007a/2013-A01450-45].

2.2.2 Stimuli

Stimuli were programmed and generated using MATLAB R2013 (http://www.mathworks.com/). The sound stimuli were sequences of normally-distributed, 16-bit pseudo-random numbers with a zero mean, which were played at a sampling frequency of 44.1 KHz. To ensure that the sounds are different every time, we reset the seed of the pseudorandom number generator of MATLAB on every trial. We

constructed Cyclic (*CN*) and Non-Cyclic (*N*) stimuli, both lasting 1 s in duration (Audio samples can be found at http://m4.ups-tlse.fr/). A CN was generated as a 500-ms pseudo-random segment of sound that was presented twice back to back (cycled). An N was generated as a 1000-ms pseudo-random segment. The spectrograms of such Gaussian white noises are flat, with no distinctive variations in frequency over time. Therefore, to illustrate the cyclic nature of these sounds, we plotted the actual amplitude variations over time (Figure 2.2 A). This shows that the amplitude variations in the first and second halves are identical in a CN but not in an N. Over the experiment, participants were presented with 4 variations of the CNs (explained below) while all the Ns were uniquely generated and heard only once. The generation and exemplar amplitude variations in modified CNs are illustrated in Figure 2.2 (B and C).

Target CN: This was a uniquely generated CN that was presented several times to the participant over the learning and testing sessions of the experiment.

CN: This was a uniquely generated CN that was heard only once throughout the experiment.

Looped target CN: This was a modified version of a target CN. For looping a target CN, a random time point was chosen from its first half, the sequence was cut at this point and the preceding segment was pasted at the end.

Scrambled target CN: A modified version of a target CN was created by segmenting the first half (500ms) into several bins of equal size, which were randomly shuffled and then played back to back to create a CN.

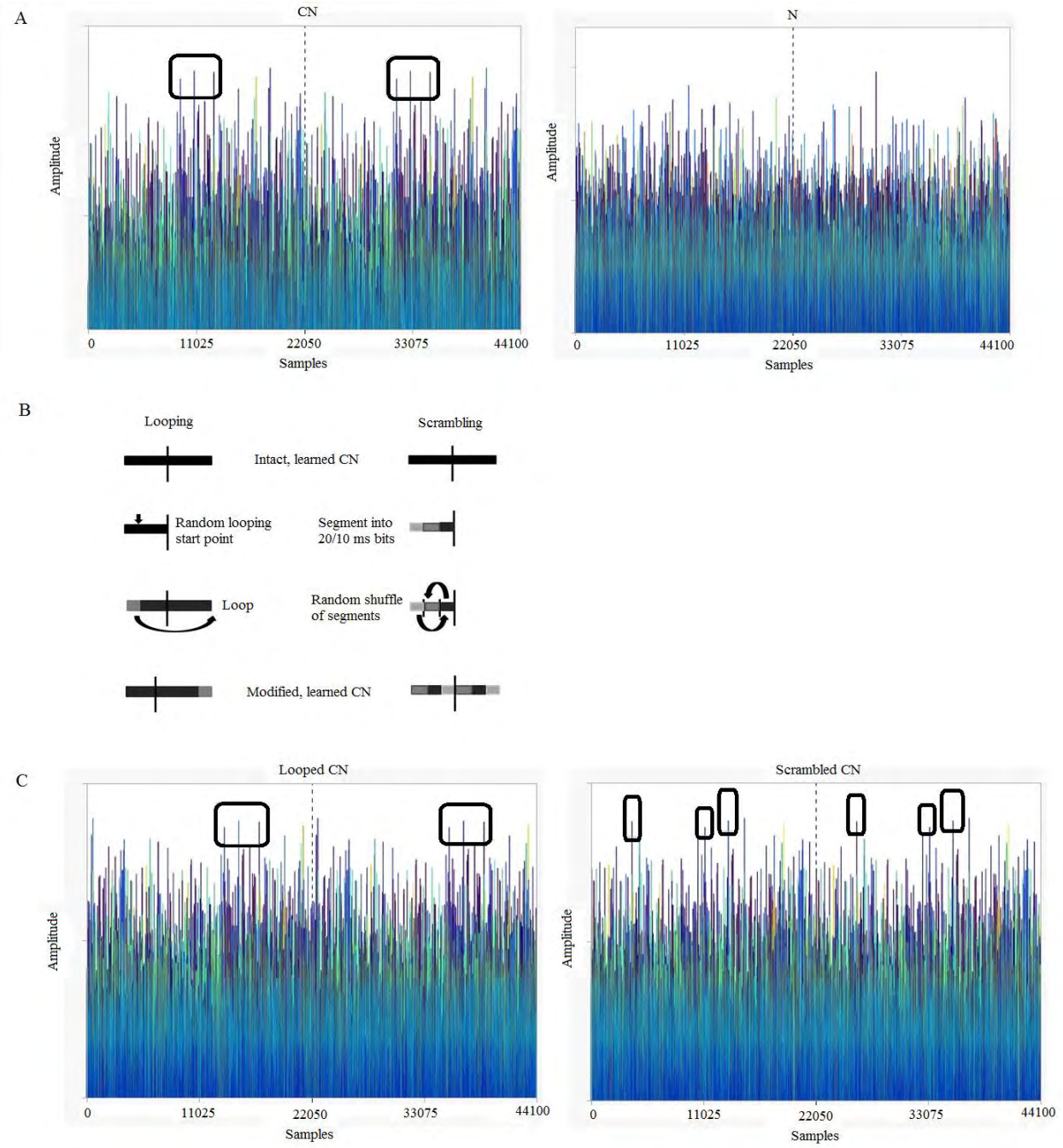


Figure 2.2: Exemplars of 1-s Gaussian white noises (sampling frequency = 44.1 kHz) and acoustic transformations used in the experiment. a) Cyclic noise (CN) vs. non-cyclic noise (N): Gaussian noises typically show small amplitude variations over time. The first and second halves of a CN are identical, while an N is completely random. b) Transformations used to loop and scramble the learned CNs in the testing session. For looping, a random time point was chosen in the first half of the sound and the sound portion preceding this time point was shifted to the end. For scrambling, the first half of the cyclic sound was cut into segments of 20 ms for version 1 and 10 ms for version 2, the segments were randomly shuffled and the resulting 500-ms sound was played back to back to create a scrambled CN. c) Looped and Scrambled sounds: amplitude variations over time of exemplar looped and scrambled (20 ms) versions of the CN shown in a). The color scheme of figures 1a and 1c is graded as a function of sound amplitudes, in order to facilitate identification of repeating features.

It is important to note that each presentation of a looped or scrambled CN was different to prevent learning of one exemplar of the looped/scrambled version of the target CN throughout the session. Looped and scrambled CNs were presented to the participants only during the testing session.

To further understand how scrambling and looping affect the acoustic properties of a CN, we calculated Fourier transforms of an exemplar CN and its variants. Variants were created similar to the looped and scrambled sounds. Bin sizes of 250 ms, 100 ms, 50 ms, 20 ms and 10 ms were used to create 5 distinct scrambled versions of the exemplar CN. The difference in amplitude between spectra of these variants and spectrum of the original CN is plotted in Figure 2.3, with frequency bins (10 samples/bin) ranging from lower bands to higher bands on the X axis. While looping and 250-ms scrambling does not change the amplitude spectrum at any frequency, scrambling using 100-ms or smaller bins affects the amplitude spectrum at all frequencies.

2.2.3 Task

All participants performed 2 sessions of a forced-choice discrimination task, 4 weeks apart. Each trial started with participants hearing a Gaussian noise of 1 second, after which they had to discriminate the sound as cyclic/non-cyclic. Participants did not receive any feedback about their performance. All trials were presented in a randomized order. After session 1, each participant's performance was analyzed (as explained in the analysis section) and learned target CNs were selected for the session 2.

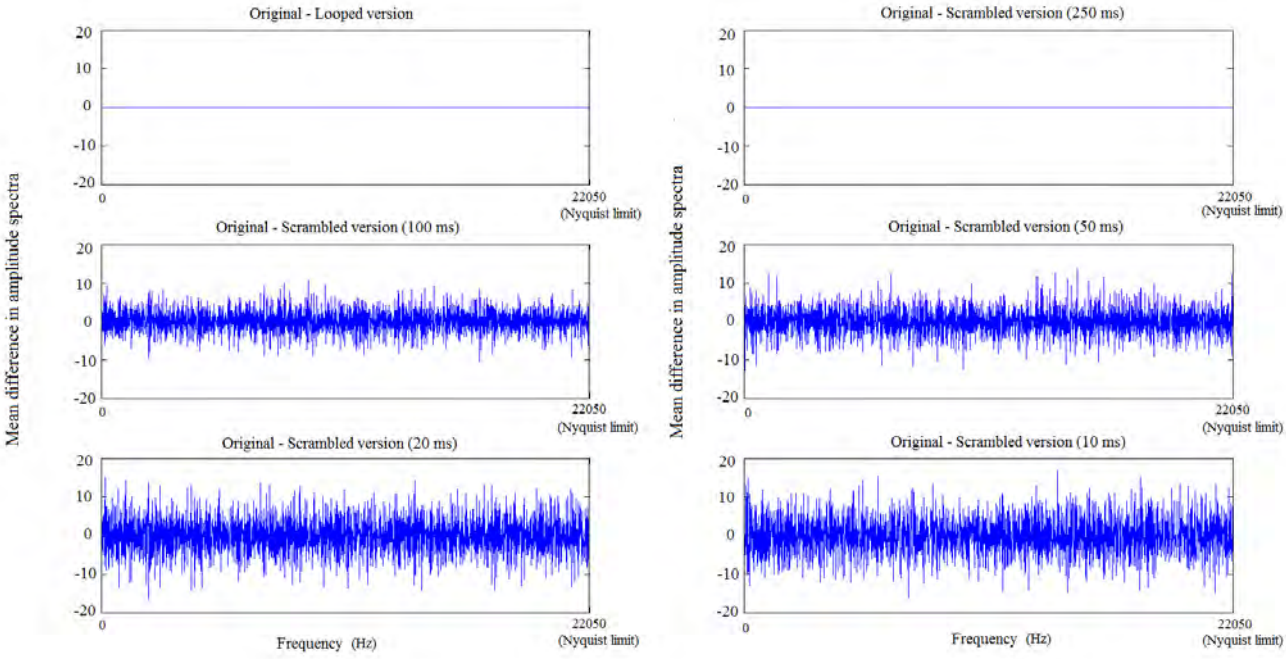


Figure 2.3: Changes in frequency features - in low, mid and high frequency bands - of a CN due to looping and scrambling with increasing bin sizes. The maximal frequency on the X axis corresponds to the Nyquist frequency (22050 Hz) and the spectrum amplitude *difference* between original and looped/scrambled versions of a CN is plotted on the Y axis. With decreasing bin size, the difference between the resulting scrambled sound and the original sound increases, leading to greater difference in amplitude spectrum from the original, across all the frequency bands.

Participants on average took about an hour to complete each experimental session, not including training. The training session took on average 15 minutes. We provided participants with scheduled breaks between blocks and participants were informed that they could also pause within a block to take breaks as necessary.

2.2.4 Procedure

Both experimental sessions included 10 blocks of 80 trials each. The first session included a training part followed by an implicit learning part. The second session was the testing part.

Each participant was assigned to perform one of 2 versions of the study. Differences between both versions are explained below.

Session 1: Training

Before starting the learning part of the experiment, all participants underwent a training session, during which they listened to CNs and Ns of varying durations. This training was intended to habituate participants to detect cyclic patterns in random noise. Each training stage was repeated with new sounds until participant reached performance criterion. To explain the difference between cyclic and non-cyclic sounds, participants first listened to samples of 5-s cyclic sounds constructed as 10 repeats of a 500 ms random noise segment and 5 second non-cyclic random noise sounds until they could confidently differentiate between the two types of sounds verbally. Participants started the training phase by listening to 5 CNs (5s, 10 repeats of 500 ms segment) and 5 Ns (5s), in random order (training stage 1). After each sound was presented, participants had to indicate via a keyboard button press if the sound was cyclic or not. They were then given feedback about their response. Once they had correctly identified all CNs they moved to the stage 2, during which they were presented with 20 CNs (2s, 4 repeats of a 500 ms segment) and 20 Ns (2s). Once participants achieved a global accuracy of 80% of correct responses for the CNs they moved to the stage 3 and were presented with 20 CNs (1.5s, 3 repeats of a 500 ms segment) and 20 Ns (1.5s) until they achieved a global accuracy of 70% of the CNs. At any stage of the training, participants who did not reach criterion ended their participation in the study. The training was identical for both versions of the experiment.

Session 1: Learning

Participants performed 10 blocks of the forced-choice discrimination task (as described earlier) immediately after training. In each block, participants were presented with 40 *Ns*, 20 *CNs*, and 20 repeats of a unique *target CN.*

Version 1

Each participant was randomly assigned to 1 of 2 possible sets of 10 target CNs. This was aimed at testing the existence of any systematic biases to detect cyclicity in some target CNs over others.

Version 2

All participants heard the same set of 10 target CNs (set 2 of version 1).

Session 2: Testing

As far as the participant was concerned, this session consisted in an identical forced-choice discrimination task, similar to the one performed in the learning session. However, the stimuli used were different: for each participant we created a list of the "learned" target CNs (discrimination performance of at least 80%, i.e., a sound which participants correctly discriminated as cyclic at least 16 out of the 20 times they heard it), which was further used to create looped and scrambled CNs. In addition to Ns, CNs and Target CNs, participants were presented with looped CNs during 5 blocks and with scrambled CNs during the other 5 blocks (block order was randomized). Each block included 40 *Ns*, 10 *CNs*, 10 *target CNs* (chosen randomly on each trial from a list of

learned target CNs for each participant) and 20 *modified (looped/scrambled) target CNs*. The scrambled sounds that participants heard were different based on the version they had been assigned to, as explained below.

Version 1

Participants assigned to version 1 of the experiment were presented with learned target CNs scrambled using 20-ms time bins. That is, the first half of a learned target CN was cut into 25 bins of 20 ms (882 samples in each bin) before shuffling to create a scrambled target CN.

Version 2

Participants assigned to version 2 of the experiment were presented with learned target CNs scrambled into 10-ms time bins. That is, the first half of a learned target CN was cut into 50 bins of 10 ms (441 samples in each bin) before shuffling to create a scrambled target CN.

Finally, we were interested in analyzing how two parameters – sleep and sound imagery – might influence learning and memory in our paradigm. Numerous studies have shown the influence of quality of sleep (review, (Walker and Stickgold, 2014)) in learning and memory for different types of stimuli. To assess quality of sleep, participants maintained a sleep diary, similar to those used previously (Mary et al., 2013), during the 4 weeks between learning and testing sessions. During the testing session, Participants also filled out St. Mary's sleep questionnaire (Ellis et al., 1981) regarding their last night's sleep quality. Lastly, to assess the influence of sound

imagery on the ability to do the discrimination task, participants also filled out a sound imagery questionnaire (Willander and Baraldi, 2010).

2.2.5 Analysis

Analysis was done using MATLAB and statistical tests were performed using JMP (Version *12.* SAS Institute Inc., Cary, NC, 1989-2007).

Learning session analysis:

The proportion of hits and false alarms in each block was calculated, for all participants. The correct identification of a CN (CNs and target CNs) was considered a hit and the incorrect identification of an N as a cyclic noise was considered a false alarm. All target CNs that were correctly identified in at least 80% of the trials were considered *learned target CNs.* The list of *learned target CNs* was subsequently used to create the testing session stimuli for each participant.

Moreover, we investigated any systematic biases in detecting cyclicity in some target CNs over others. For each target CN presented in the learning session, the proportion of participants who actually learned the sound was determined. This was done for each set of target CNs over the two versions of the experiment.

Individual discrimination performance was computed over the 10 blocks, using the principles of signal detection theory. We calculated individual *a'*, a non-parametric measure of participants' sensitivity to differences between signal (target) and noise (distractor), i.e. cyclic vs. non-cyclic stimuli (Pollack and Norman, 1964; Stanislaw and

Todorov, 1999). While sensitivity is traditionally evaluated using d', an assumption for using d' is that signal and noise distributions have equal standard deviations. In our experiment, since the signal trials include different subtypes of trials (CNs or target CNs) but not the noise trials (Ns), a' is a better estimate for sensitivity than d'. A' was calculated using the formula provided by Stanislaw and Todorov:

$$A' = 0.5 + \left[\text{sign}(H - F)(H - F)^2 + \frac{|H-F|}{4*\max(H,F)-4*HF}\right] \qquad (1)$$

$$\text{Sign}(H - F) = +1 \text{ if } (H - F) > 0 \text{ and } -1 \text{ if } (H - F) < 0$$

Where H = proportion of Hits for the signal trials and F = proportion of False Alarms for distractor trials.

Any participant with a' < 0.5 was excluded from the analysis (and from subsequent participation in the testing session) since this implied that this participant's performance was at chance level.

Testing session analysis:

The proportion of hits (correct identification of *learned target, looped, scrambled* and *novel CNs* as cyclic) and false alarms (incorrect identification of *Ns* as cyclic) was calculated for each individual. The discrimination rate for CNs was determined individually as the number of times (out of 20 presentations within a block) a CN was correctly discriminated.

To differentiate between participants who were merely good at detecting noise cyclicity from those demonstrating a preferential bias towards previously learned cyclic sounds, i.e. an implicit memory effect, we compared discrimination rates for *learned target* and *novel CN* trials in each individual. A participant who had truly learned a *target CN* would more accurately detect cyclicity for this noise over a *novel CN*. To ensure that any observed preferential bias to discriminate learned target CNs was not due to within-session rapid learning, the discrimination rate for learned target CNs was also analyzed as a function of time.

Moreover, to investigate the relationship between how well a sound was learned, quantified as discrimination performance in the learning session, and subsequent memory resistance to acoustic transformations, we compared a' during learning ($a'^{learning}$) with discrimination rate for *intact learned, looped* and *scrambled target CNs* in the testing session. A high $a'^{learning}$ would mean participants accurately detect cyclicity during the learning session.

Lastly, scores from the questionnaires on sleep quality and sound imagery were correlated with a' values across our group of participants.

2.3 Results

Based on individual performances in the training and learning session [inclusion criteria: a' > 0.5], data from 16 (of 26) participants in version 1 and data from 9 (of 11) participants in version 2 were included in the analyses. We first looked at the training performance and the number of times a participant performed each training stage during the learning session had no effect on the discrimination performance during the testing session (a' testing): training stage 1 [$F(1,25) = 0.09$, $p = 0.75$], training stage 2 [$F(1,25) = 1.98$, $p = 0.17$] and training stage 3 [$F(1,25) = 0.03$, $p = 0.86$] (Figure 2.4).

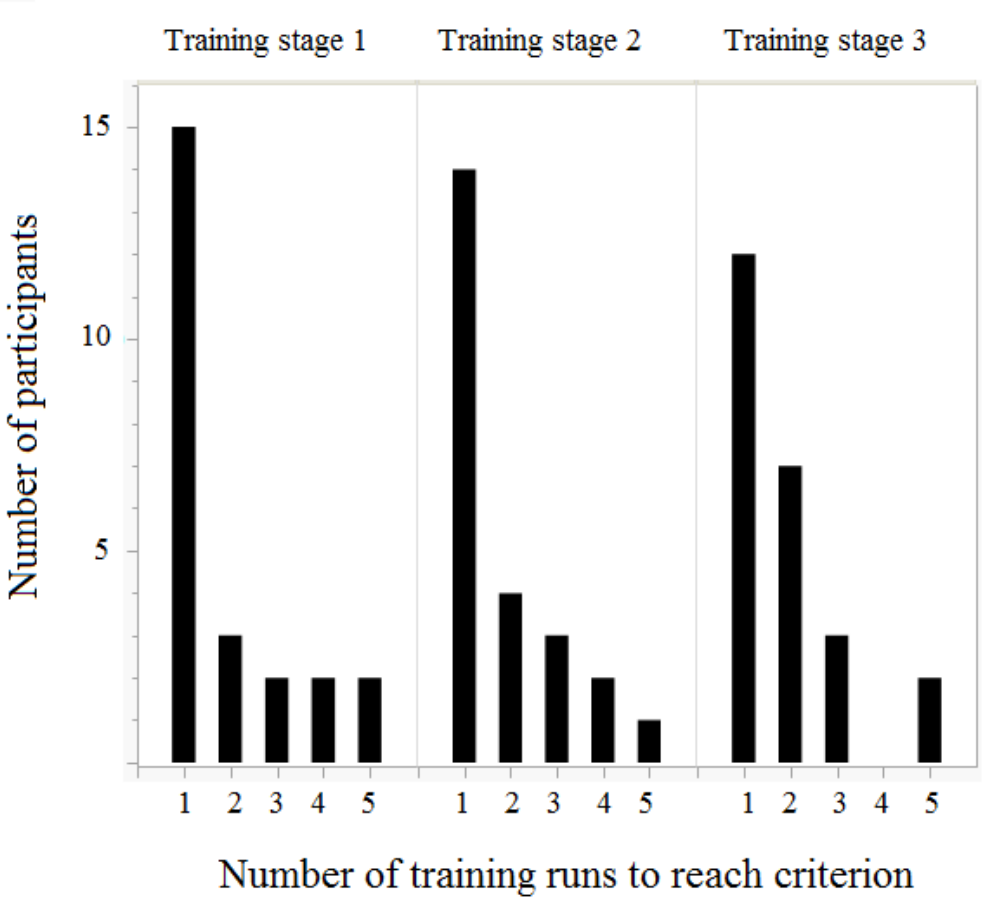


Figure 2.4: Performance during the training session, representing number of runs to reach criterion before moving onto the next stage. Participants who did not reach criterion within 5 runs of any stage discontinued the experiment.

Since the learning session followed an identical procedure and resulted in equivalent discrimination sensitivity ($a'^{learning}$) in both versions [$F(1, 25) = 0.4287$, $p = 0.52$], data from the first session for all 25 participants were pooled. Data from the training session for both versions is summarized in figure 2.4. Individual a' values ranged between 0.53 and 0.97 (mean a' = 0.73, S.D = 0.71). The number of sounds learned by participants within each set of target CNs was computed, showing no preferential bias for some sounds over others in set 1 [$F(9,80) = 0.26$. $p = 0.98$] and set 2 [$F(9,170) = 1.09$, $p = 0.37$] (Figure 2.5). Moreover, the proportion of participants who learned the target CNs did not differ between both sets [$F(9, 20) = 0.24$, $p = 0.98$] (Figure 2.5).

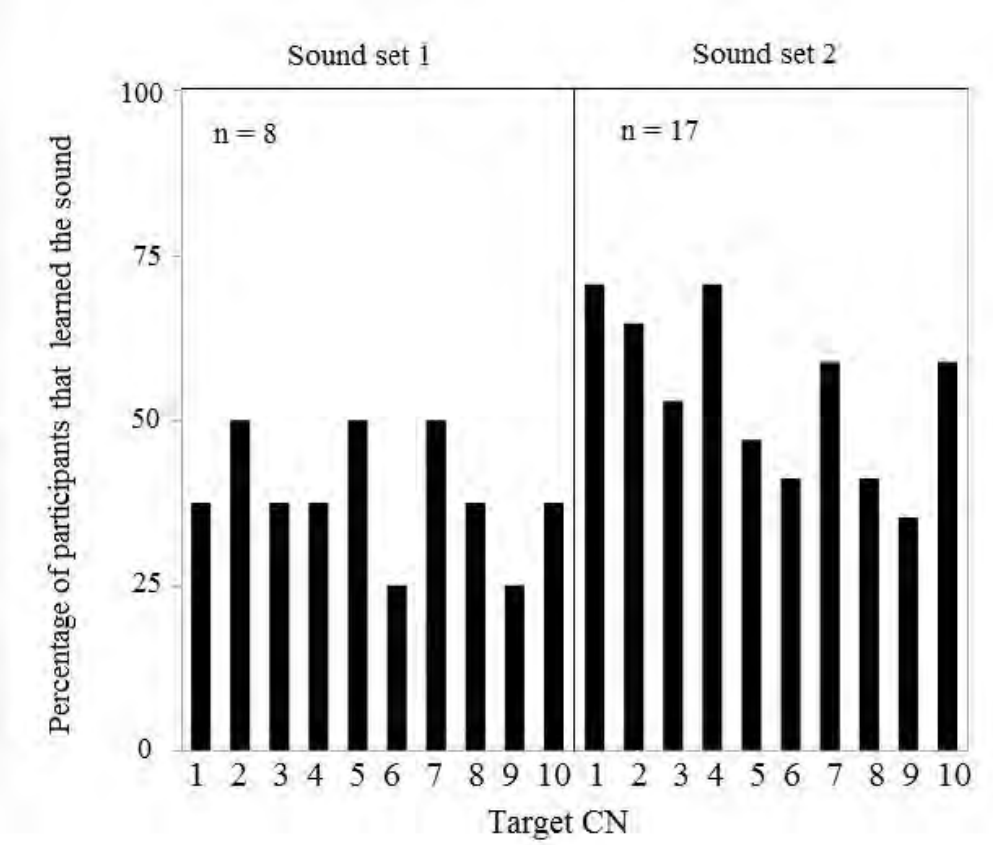
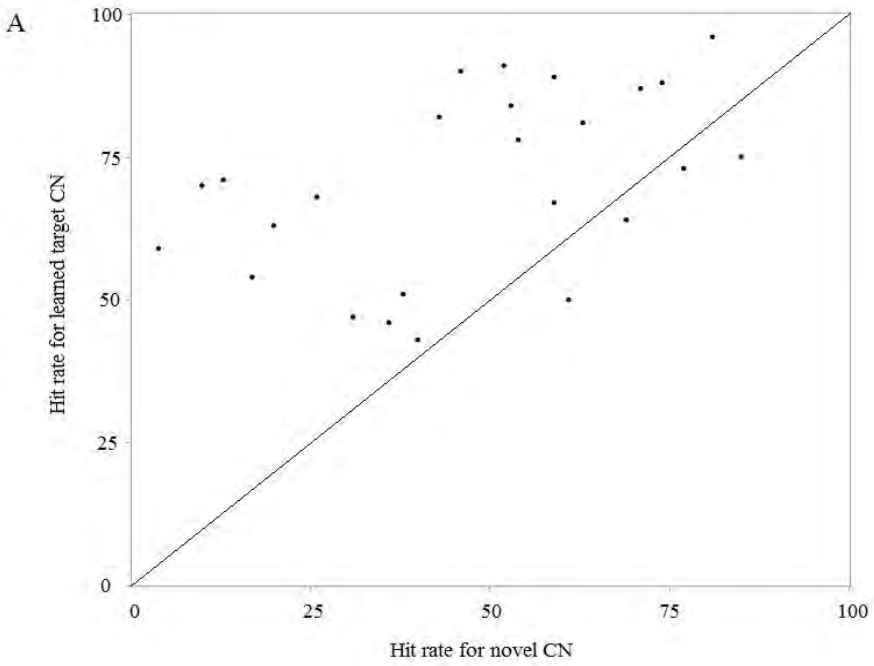
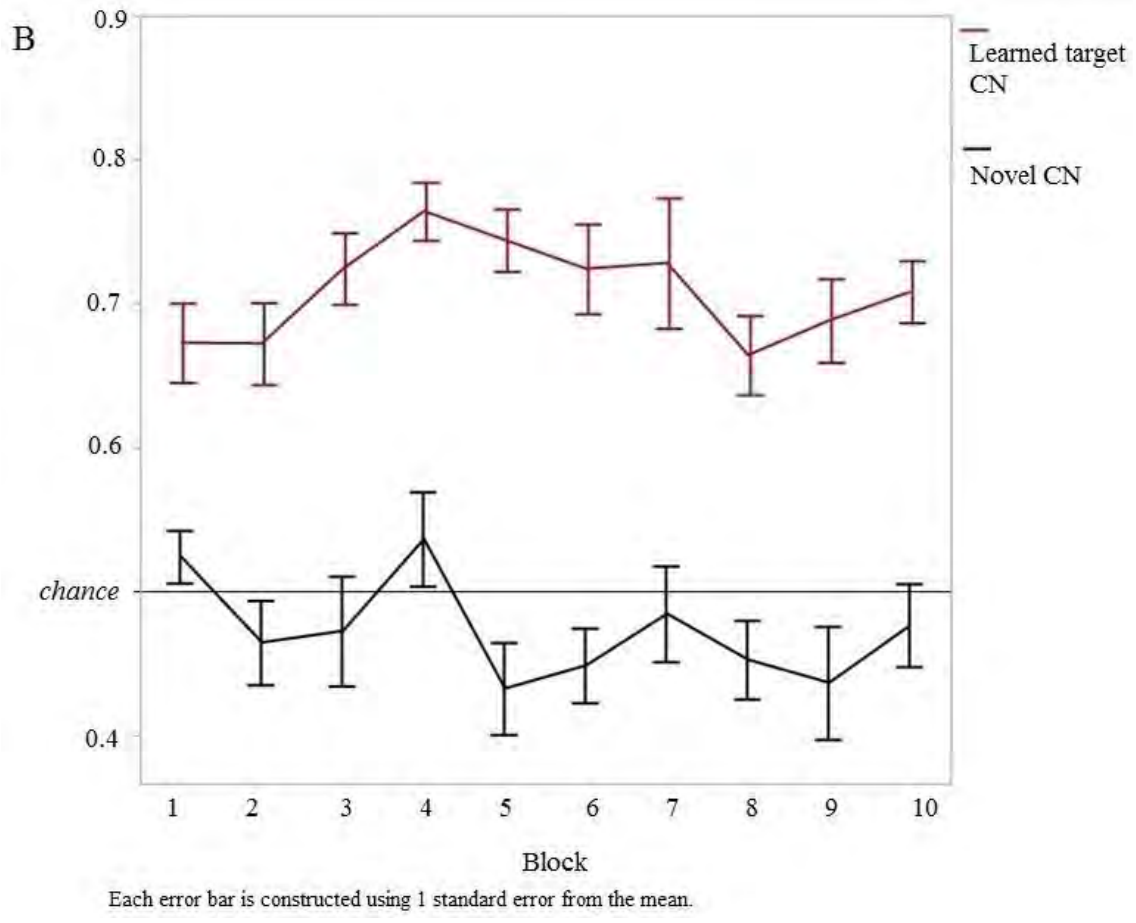


Figure 2.5: Learning session results for both sets of 10 target CNs: Each target CN was learned by a variable percentage of participants, i.e., no target CNs were systematically learned by all participants.

Participants performed the testing session a month after the learning session (mean interval 30.96±4.2 days, range 23 - 41 days). In the testing session, discrimination sensitivity ($a'^{testing}$) again did not vary between versions 1 and 2 of the experiment [$F(1,25) = 0.0057$, $p = 0.94$] and therefore these data were pooled (n=25). Detection of cyclicity in novel CNs did not change across the two sessions (Mean difference = 0.58, S.E = 3.8, $t(25) = 0.15$, $p = 0.8803$) indicating that participants' performance in the task was similar over the four weeks. Within the testing session, discrimination rates were significantly higher for learned target CNs compared to novel CNs [$F(1,25) = 7.03$, $p< 0.014$] (Figure 2.6 a). This suggests that participants had memory for the CNs previously learned in the first session. Furthermore, to ensure that this higher discrimination rate for learned target CNs did not result from learning of features throughout the testing session (as opposed to long-term memory for features from the first session), the evolution of discrimination rates for learned vs. new CNs was analyzed over time. A two-way repeated-measures ANOVA on discrimination rates was computed, testing main effects and interaction of within-subjects factors 'trial type' (2 levels, 'learned target CN' and 'novel CN') and 'block' (10 levels). Trial type was the only significant predictor of performance [$F(1,200) = 313.696$, $p<0.0001$] irrespective of block [$F(9,200) = 1.57$, $p=0.127$]. The effect of trial type was equivalent across blocks [$F(9,200) = 1.06$, $p=0.394$]. These results were confirmed by the absence of correlation between hit rate for learned and novel CNs over the 10 blocks [$F(1,100) = 0.14$, $p = 0.71$; $R^2 = 0.001$, slope = -0.03, intercept = 0.7, $p = 0.71$]. These results are shown in Figure 2.6 b. Progression of the other trial types, i.e. looped and scrambled learned CNs, are summarized in figure 2.6 c.

2.6





Each error bar is constructed using 1 standard error from the mean.

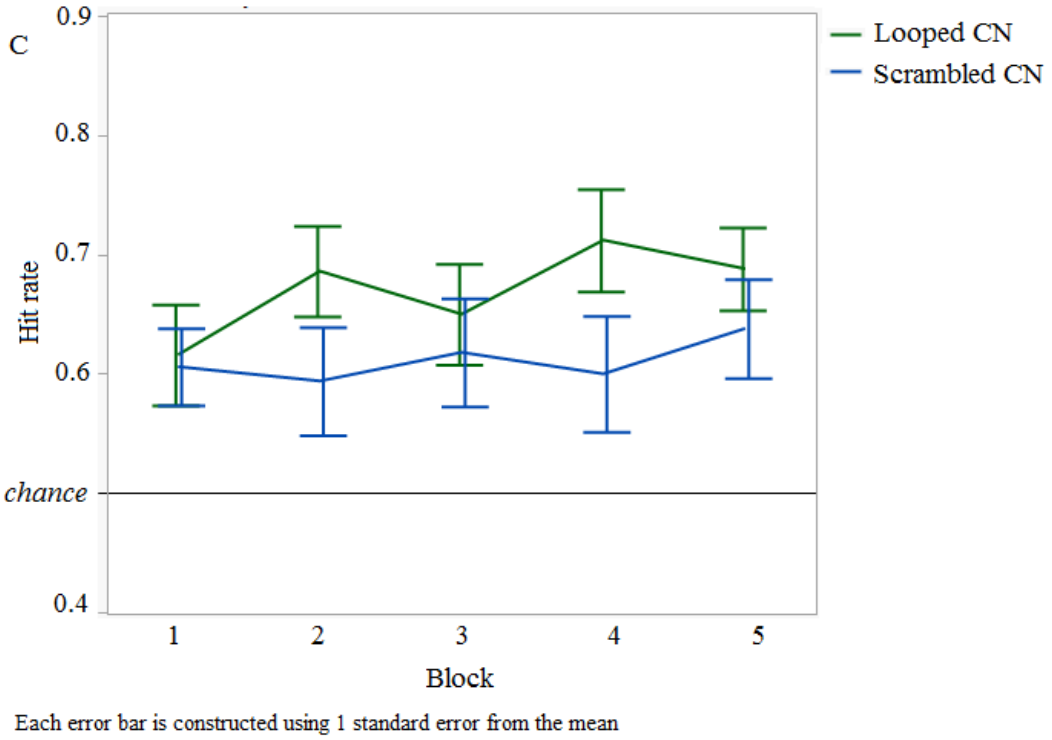


Figure 2.6: Discrimination performance for learned target CNs vs. novel CNs in the testing session: a) Relationship between discrimination rates of learned target and novel CNs in each participant. Participants above the diagonal show higher rates for learned vs. novel CNs, suggesting that memory facilitated the discrimination task. b) Discrimination rates of learned target and novel CNs over time (10 blocks). c) Discrimination rates for looped and scrambled CNs over time. Looped CNs were presented during half of the blocks, i.e. 5 blocks, and the other half included scrambled CNs. As in 2.6.b, discrimination rate for looped and scrambled trials are above chance from the first block.

Since target CNs are 'learned' within 20 presentations during the learning session, it is still possible that any observed difference in discrimination performance between learned target CNs and novel CNs is due to learning within the testing session. To investigate this possibility, we also calculated the discrimination rates for target and novel CNs during the first 10 trials of the first experimental block of the testing session. Repeated measures ANOVA revealed that mean discrimination rate for intact target

CNs during the first 10 trials of first block of the testing session was significantly higher than mean discrimination rate for novel CNs [$F(1,20) = 20.098$, $p = 0.0003$]. This finding indicates that participants truly demonstrated recognition memory for target CNs.

Since participants had long-term memory for learned target CNs, we further analyzed discrimination rates for all types of CNs in the testing session. Since scrambled target CNs were different in versions 1 and 2 of the experiment, the effect of version on discrimination rates was specifically tested. A two-way ANOVA was conducted, using within-subjects factor of 'trial type' (4 levels, 'intact target CN', 'looped target CN', 'scrambled target CN' and 'novel CN') and between-subjects factor of 'version' (2 levels). A significant effect of trial type on discrimination rates was found [$F(3,100) = 23.73$, $p<0.0001$]. There was no effect of version on discrimination rates [$F(1,100) = 2.29$. $p=0.1432$], and no interaction between both factors [$F(3,100) = 0.99$, $p = 0.4$]. Since there was no evidence for any effect of version or interaction, data were pooled across both versions (Figure 2.7a) and differences between trial types were further examined. Tukey's Honestly Significant Difference (Tukey's HSD) tests showed that discrimination rates for novel CNs were lower than discrimination rates for all other CNs; that is, detection in intact target [effect size ($mean_{(i)}$-$mean_{(j)}$) = 24.6, $CI_{95\%}$ = (16.5 , 32.7), $p = 0.0001$], looped [effect size = 19.23, $CI_{95\%}$ = (11.1 , 27.3), $p = 0.0032$] and scrambled [effect size = 15.67, $CI_{95\%}$ = (7.6 , 23.7), $p = 0.0362$] CNs were all significantly higher than novel CNs. Discrimination rates for intact target and looped trials were equivalent [$p = 0.8004$], and so were the discrimination rates for looped and scrambled trials [$p = 0.8464$]. However, discrimination rates for learned intact trials were higher than for scrambled trials [effect size = 8.93, SE = 3.07, $p = 0.3164$]. We were also

interested in any performance difference for the scrambled trials between versions 1 and 2, since the bin sizes were different in the two versions. As shown from the two-way ANOVA, discrimination rates were not impacted by the version of the experiment, indicating that scrambling learned CNs with bin sizes of 10 or 20 ms resulted in equivalent performance. For information, we report in Figure 2.7b results for scrambled CNs, for the 2 versions separately.

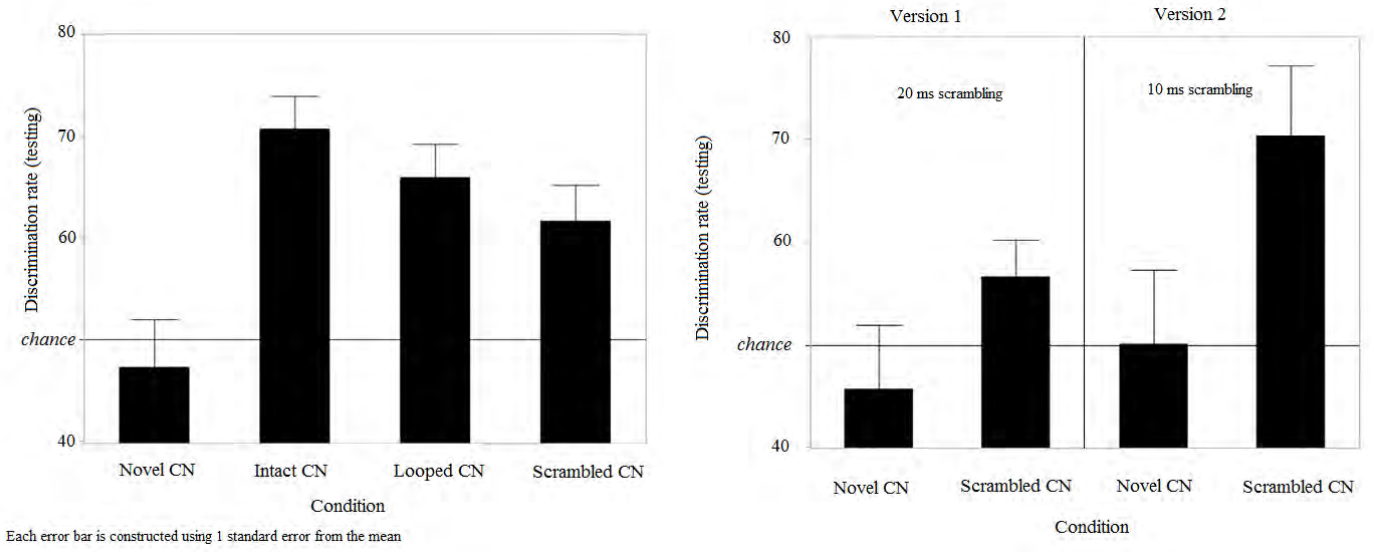


Figure 2.7: Discrimination performance during the testing session: a) Performance for intact, looped, scrambled learned target CNs and novel CNs (n=25) and b) Discrimination performance for scrambled trials with 20 ms and 10 ms bin sizes (n = 16 in version 1 and n = 9 in version 2).

Individual discrimination performances in the testing phase as a function of discrimination efficiencies (quantified as a') during the learning phase for learned CNs and modified forms of the learned CNs were examined. Linear regression between hit

rate$_{(testing)}$ and a'$_{(learning)}$ showed that a' during learning did predict later detection of cyclicity in looped [R^2 = 0.272, slope = 81.95, intercept = 5.71, p = 0.0075] and scrambled CNs [R^2 = 0.366, slope = 102.59, intercept = -13.73, p = 0.0013]. The correlation between a' during learning and discrimination rate for intact learned CNs [R^2 = 0.153, slope = 60.36, intercept = 26.32, p = 0.053] was just below significance. These results show that a' during learning was a significant predictor of further accuracy to discriminate modified versions of learned CNs 4 weeks later. This is shown in Figure 2.8.

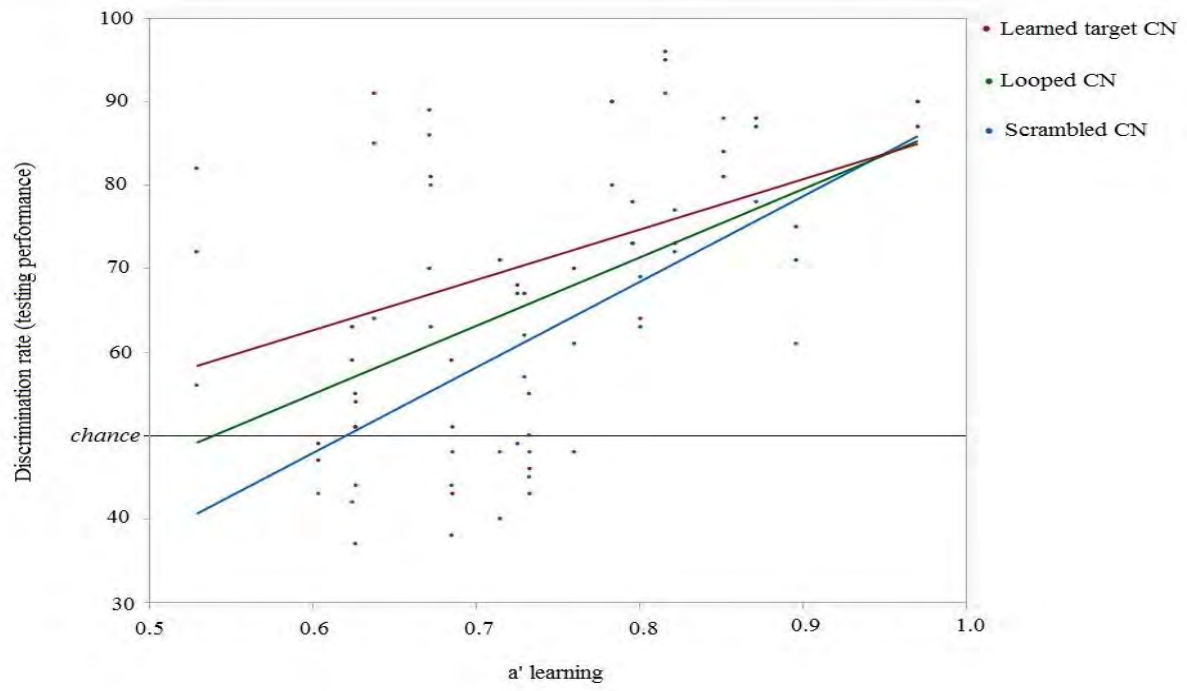


Figure 2.8: Relationship between discrimination rates of CNs in the testing session and learning efficiency (represented as a') for all participants (n = 25).

Lastly, the relationship between discrimination performance in the learning and testing sessions and parameters quantified using sleep and sound imagery questionnaires was investigated. The results are summarized in figure 2.9. There was no correlation

between discrimination ability in the testing phase and self-reported sleep quantity [R^2 = 0.02, slope = 0.007, intercept = 0.7, p = 0.55] or quality of sleep [R^2 = 0.004, slope = -0.002, intercept = 0.79, p = 0.76] the night before testing. There was no correlation between discrimination rates and alertness the day of testing [R^2 = 0.05, slope = 0.02, intercept = 0.69, p = 0.29]. We also found no link between sound imagery scores, as assessed by St Mary's questionnaire, and discrimination in the learning [R^2 = 0.07, slope = -0.03, intercept = 0.85, p = 0.32] and testing [R^2 = 0.09, slope = -0.04, intercept = 0.89, p = 0.23] sessions.

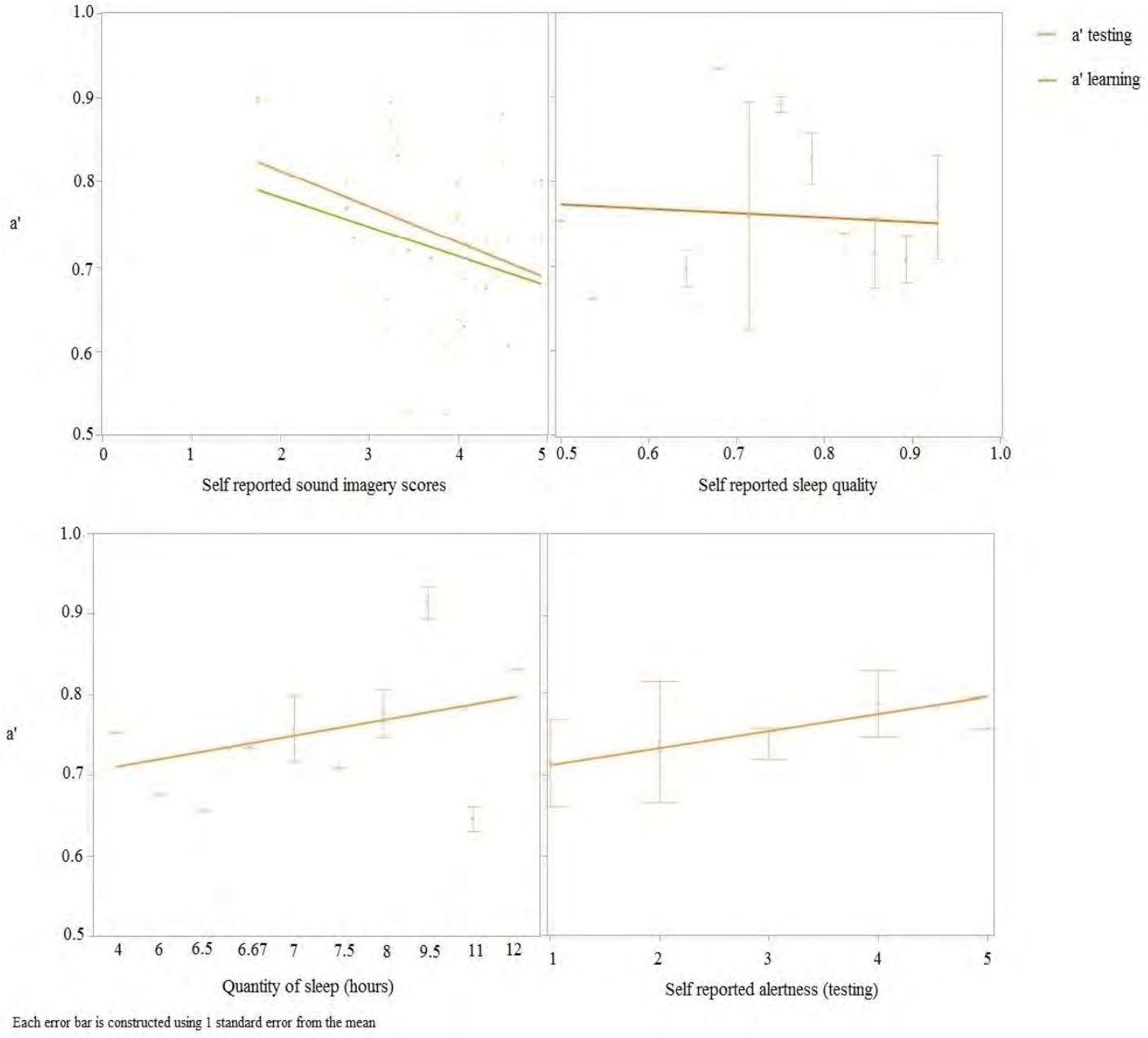


Figure 2.9: Correlations between sleep quality, sound imagery and discrimination rates of CNs (measured as a'). Clockwise from the top-left: A: Correlation between sound imagery (measured using the French version of (Willander and Baraldi, 2010)) to learning and testing performance. B: Correlation between self-reported sleep quality (measured from a subset of questions from the St. Mary's sleep questionnaire) and testing performance. C: Positive correlation between self-reported alertness the day of the testing session (measured from another subset of questions from the St. Mary's sleep questionnaire) and testing performance. D: Positive correlation between self-reported sleep quantity (measured from a third subset of questions from the St. Mary's sleep questionnaire) and testing performance. Overall, none of the parameters we measured significantly influenced discrimination performance.

2.4 Discussion

The present results confirm our hypothesis that features in meaningless sounds can be learned and retained over several weeks. The results also demonstrate the *robustness* of this memory to acoustic transformations: despite a decrease in the preferential bias to detect learned features with increasing degree of transformation from the original, participants were more accurate to detect cyclicity in highly degraded versions of learned sounds in comparison to novel cyclic sounds.

The *quality of learning* was a predictor of this memory to survive acoustic transformation. While models of sleep and memory predict that stored features are subject to opposing factors that selectively strengthen (reactivations during sleep) (Rudoy, Voss, Westerberg & Paller, 2009) or weaken (internal pruning based on probability estimations of re-occurrence) (Kim et al., 2014) the memory trace, sleep parameters quantified by self-report measures did not correlate with discrimination performance; more objective measures of sleep are necessary to understand the role of sleep in memory for Gaussian sounds.

Regarding robustness of memory to acoustic transformations, we found that participants had equivalent implicit recognition memory for intact and looped (onset-shifted) versions of a learned sound, clearly demonstrating that feature learning was not restricted to sound onset. Instead, learned acoustic features that facilitate implicit recognition may be scattered throughout the sound. We also surprisingly observed long-term implicit recognition of scrambled versions of learned sounds, where only

small bin sizes of 10 and 20 ms were preserved and temporal context of learned features was lost. Although the limit of memory capacity has been discussed since Miller proposed the concept (Miller, 1956), few studies have investigated the capacity limits of implicitly-encoded purely sensory memory. The upper capacity limit in working memory for Gaussian noises was found to be around 100 ms for individual spectro-temporal features, while the lower resolution limit was unclear (Kaernbach, 1993; Kaernbach, 2004). Our results demonstrate that this lower resolution limit could be as short as 10 or 20 ms. Since each presentation of a scrambled sound was randomly generated, new features greater than 10ms in length could not be learned throughout testing. As shown in our analysis, scrambling modifies the spectral features of original sounds as a function of bin size, with these modifications staying nearly uniform across higher and lower frequency bands. This observation renders the coding of sound frequency features an unlikely mechanism to explain long-term implicit recognition. An alternative explanation would be that participants were able to store temporal features shorter than 10ms. Interestingly, participants who accurately discriminated cyclic and non-cyclic sounds during the learning session also had higher implicit recognition memory for looped and scrambled versions of learned CNs, suggesting that the size of a stored feature is inversely proportional to encoding efficiency.

Our data also demonstrate that these learned features vary between participants, and that no single feature could be learned by all participants. The phenomenon of stochastic resonance, where *optimal* noise can enhance the periodicity of a weak signal causing the signal to rise above the threshold for detection (Wiesenfeld and Moss, 1995), as discussed in the introduction section 1.3B, puts this finding into perspective. Based in

this idea, we speculate that specific acoustic features (weak signal) of a Gaussian sound may be preferentially enhanced when added with baseline neural activity (optimal noise) for a given individual, resulting in different features being encoded by different participants. Further support for this hypothesized mechanism of learning features in meaningless stimuli comes from the MEG study using a similar discrimination task conducted by Luo and colleagues (Luo et al., 2013), discussed in the introduction section 1.5A**.** The authors found that the phase of auditory cortical neural responses change and track learning of target CNs in the theta (3-8 Hz) range. They also demonstrate that different learned target CNs induce diverse phase pattern responses (figure 1.22). These results suggest that as features in target CNs are learned, phase-mediated temporal encoding specific to the learned feature occurs in the auditory cortex. Since white noise (which doesn't contain acoustic features or edges) does not reset the phase of ongoing oscillations (Luo and Poeppel, 2012), stochastic resonance could contribute to feature detection. These ideas raise interesting hypotheses for future testing. Consistent with the results obtained by Andrillon and colleagues (Andrillon et al., 2015), our results suggest that individual factors influencing neural activity, such as attention, impact the encoding of acoustic features.

Our results are not in line with traditional models of sensory memory (discussed in the introduction section 1.3A**.** As argued by Winkler and Cowan, our data suggest a need for a better model to explain mechanisms of auditory sensory memory. Data from our scrambling condition show that there is long-term memory for purely sensory features. Additionally, *a'* during learning influenced the robustness of implicit recognition memory. Fluctuations in the attentional network as well as bottom-up, feature-based

attention invoked by the stimuli may have affected participants' accuracy to differentiate cyclic and non-cyclic sounds during learning and memory formation, therefore challenging the first claim regarding sensory memory, that it "forms independently of attention" [the claims of the existing models of sensory memory as summarized by (Winkler and Cowan, 2005)]. This is further supported by individual differences in encoding a given noise feature. The fact that participants retained their preferential bias over several weeks also challenges the 4th claim of sensory memory model (short retention time). Rather, our results are more in line with predictions from the emergent memory account (EMA) (Graham et al., 2010). According to this model, the boundary between sensory perception and memory is not clearly defined and memory emerges as a result of hierarchical organization of perceptual representations that are distributed throughout the brain. Thus, this model predicts that sensory memories can rapidly form as a function of attention and number of presentations. Our results are compatible with these predictions and memory for Gaussian noise is likely the result of detecting repeating spike patterns. Attention modulates the sensory representations of these features while number of presentations influences the probability of feature detection. The models discussed by Winkler and Cowan (Winkler and Cowan, 2005) and the EMA (Graham et al., 2010) are quite different in their explanations of how sensory memory works since the EMA model is a model of all types of memory. Although our data challenge traditional models of sensory memory, further experiments specifically comparing observed (experimental data) Vs predicted (from models) features of memory need to be conducted to understand the mechanisms of sensory memory in light of these varied models of memory.

2.5 STDP model of learning CNs (executed with help from Martin Deudon).

Support for the surprising finding that short acoustic features are rapidly (Andrillon et al., 2015) and robustly (our results) stored, comes from Spike time dependent plasticity (STDP) models. The unexpectedness of our results becomes clear on comparing the observed discrimination performance (figure 2.7 a) to the predicted discrimination performance provided by the STDP model (figure 2.1) for sounds scrambled using 10-/20-ms bins. STDP models demonstrate how neurons can learn repeating spatiotemporal patterns in noise (Masquelier et al., 2008, 2009) [Discussed in introduction section 1.3B, figures 1.9 and 1.10]**.** Specifically, the finding of Masquelier and colleagues, that the first spikes were observed 4 ms after onset, provides support for the idea that very short temporal features of meaningless stimuli can be stored. Taking this further, we predicted that the Gaussian sounds used in our study would induce firing patterns in the auditory nerve similar to those observed in the afferents of this model (Masquelier et al., 2008). To test this hypothesis, we trained a similar leaky integrate and fire (LIF) neuron model on one exemplar target CN used in the experiment. The schematic of this model is represented in figure 2.10 a.

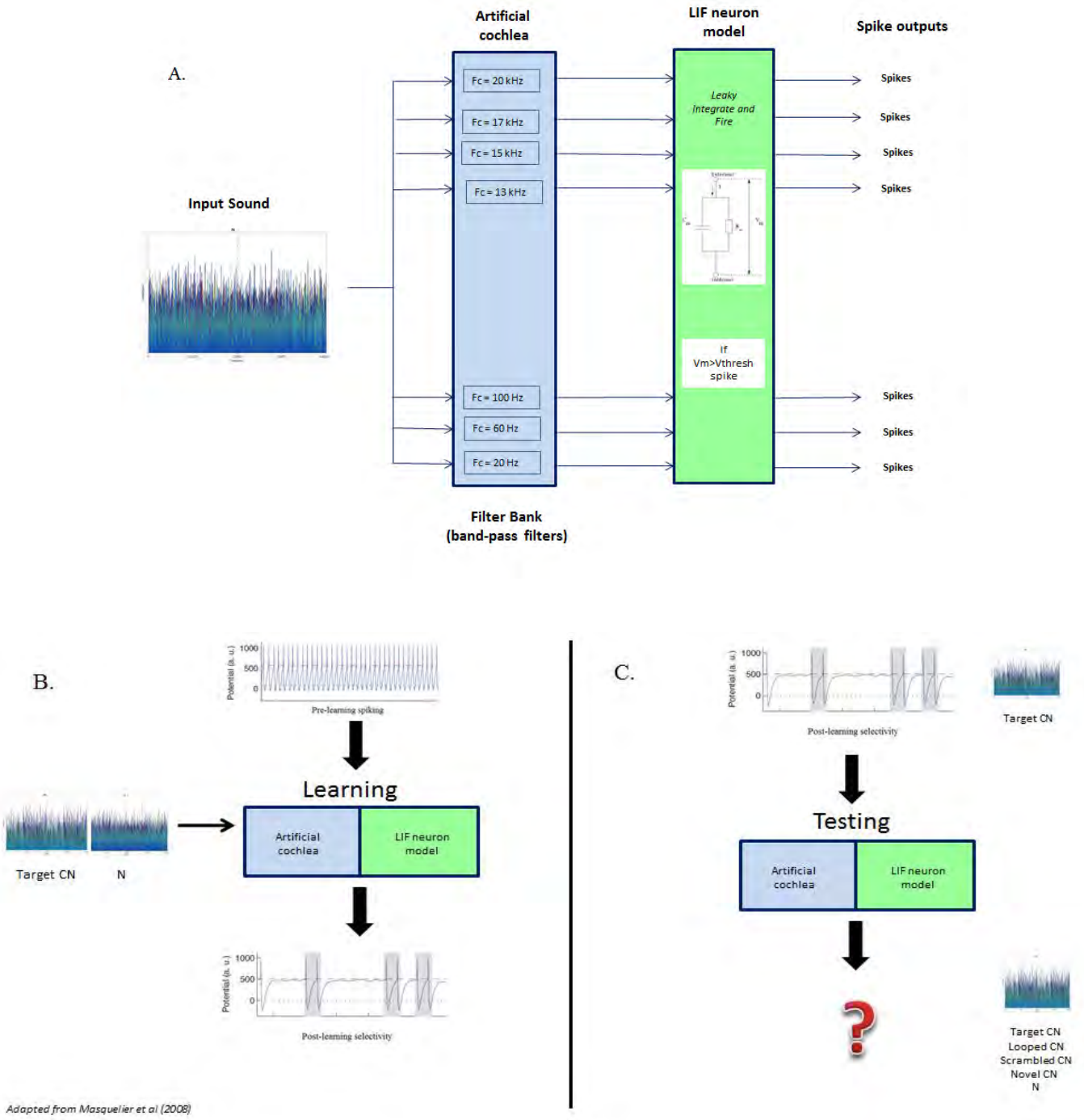


Figure 2.10: Schematic of the STDP model used to model the behavioral results. A) Representation of the model. Input sounds are first converted into spikes using a filter bank of 1000 frequency filters that effectively acts as a cochlea. 1000 outputs of this filter bank are fed into the Leaky integrate and fire neuron model previously described (Masquelier et al., 2008). B) Using this model, an exemplar target CN used in the experiment was fed into the model and the model showed selectivity of firing to target CN. C) The model was then fed with intact learned, looped and scrambled versions of the target CN as well as novel CNs and Ns to compare spike output for these conditions.

Once the model started demonstrating selectivity to this target CN, looped, scrambled and intact versions of this CN were tested using this model. The spike output of this model was then compared for all the different types of CNs (intact learned, looped, 10-ms bin scrambled and novel) as well as Ns, as represented in figure 2.10 b and c.

The spike outputs for different types of sounds are shown in figure 2.11. Within 5 presentations of the target CN, an increase in number of spikes for this target CN was observed. This was accompanied by none of the afferents spiking for a non-cyclic N. While a few of the afferents did spike for the second half of a novel CN, the number of spikes observed at any given time over all the afferents were much lower than that observed for learned target CNs. Fascinatingly, the STDP results for looped and scrambled CNs are in perfect agreement with the behavioral results observed in our experiment (figure 2.7 a). The spike output of the output afferents for looped versions of the sound is very similar to the spike output for the learned target CN, demonstrating that looping does not degrade the model LIF neuron's ability to 'recognize' a learned CN. The spike output for scrambled versions of the learned target CN lies in between the outputs observed for novel CNs and looped CNs. The model neuron only seems to 'recognize' the scrambled CN in its second half (like that of novel CNs). However, once recognized, the afferent spike outputs for scrambled versions resemble that of intact learned CNs. On increasing the scrambling bin size to 50 ms, we observed that the spike output for the scrambled versions were similar to that of learned target CNs in both halves of the CN.

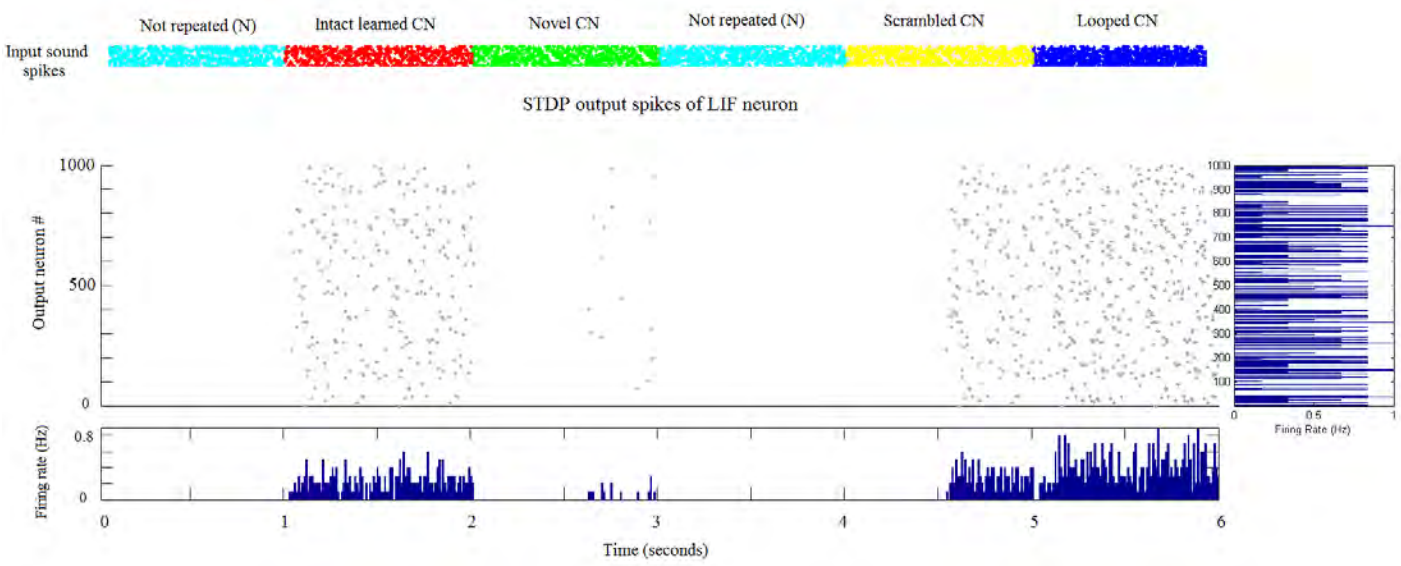


Figure 2.11 Spike outputs for afferents from a leaky integrate and fire neuron model that has specialized to respond to a target CN. When presented with degraded versions of the learned target CN, the afferents spike frequencies change as observed behaviorally.

While globally, the outputs of the LIF neuron seem to be in agreement with observed behavior (figure 2.7); does this correspondence hold up under closer scrutiny? In other words, can these results explain the nuanced behavior that participants demonstrate? One such question that arises is: if spike densities of scrambled CNs are the same as intact CNs, (at least in the second half), how is the drop in performance for scrambled CNs explained behaviorally? The answer lies, I think, in the variability observed in learning behavior. When sounds are well encoded, participants do not show a behavioral difference in detecting cyclicity in scrambled CNs and intact CNs (figure 2.8, higher a' values). If fewer patterns of a target CN are learned in by the LIF model, the spike output for the scrambled version might resemble a novel CN more than the target CN. That is, how well the model neuron learns patterns would dictate how much the

spike output of the scrambled and looped versions match that of the original target CN, thereby showing nuanced outputs similar to that observed behaviorally. Another factor to keep in mind is that STDP does not take into account confounding processes such as attentional fluctuations observed in behavioral systems. It is thus astounding that such a simple model using a single neuron using unsupervised learning can explain observed behavior so well.

These results demonstrate the eligibility of STDP as a viable biological mechanism of statistical learning and memory, at least in the auditory modality. One can imagine that a coincidence detector neuron receiving the output from a LIF neuron that has learned a target CN will spike when exposed to the learned CN, and thereby trigger recognition mechanisms downstream. When exposed to degraded versions of the learned CN, the coincidence detecting neuron would have a reduced probability of firing, the exact reduction of this probability depending on the threshold of firing and the extent to which the learned CN is degraded. This would result in downstream recognition mechanisms triggered less frequently when exposed to degraded versions of the learned CN. We can then speculate that when participants show memory for 10- and 20-ms scrambled versions of learned CNs, a few neurons that fire at/below 25 Hz acting as coincidence detectors [as discussed in the introduction section 1.3B (König et al., 1996)] probably specialized to respond to extremely short features in learned target CNs.

2.6 Conclusions and speculations

To conclude, using the frozen noise paradigm in an implicit learning protocol, we showed that participants had robust implicit recognition memory for short temporal features of meaningless sounds, and that acoustic features as short as 10 ms are possibly being stored in long-term memory. Further studies investigating the resolution limit of acoustic perception and storage are underway. To understand these results at a mechanistic level, we compared the behavioral results we obtained to results from an STDP model, a model that has been proposed to explain statistical learning of features in meaningless sounds. We demonstrated that an STDP based learning mechanism can explain the surprising finding that participants are able to recognize highly degraded versions of sounds that they have learned. We also showed that the robustness of this memory may depend on individual encoding strength.

The question of *why* we are able to store this relatively 'useless' information with such high precision is a puzzling one. Human brains have been evolving over millennia to be highly efficient. In fact, Achard and Bullmore (Achard and Bullmore, 2007) studied the efficiency and cost of brain functional networks in and fMRI paradigm. They noted that several properties such as sparse coding, parallel processing and mostly short-range connections lead to high local and global efficiency. However, storing high precision meaningless information purely on the basis of repetition, at least beyond infancy, somehow seems inefficient. So why does it happen?

A possible cause of this behavior can be understood in terms of a fundamental property of the auditory system. Unlike the visual and tactile domains, spatial precision of information is less important in audition than temporal precision. Temporal precision is essential for perceiving small differences in sound structures arriving in each ear (intra-aural time difference) which allows accurate localization of sound source. Very high temporal precision in signaling and spiking has been demonstrated right from the cochlea (Moser et al., 2006) and cochlear nucleus (Golding et al., 1995). This ability is critically important for survival of a species from an evolutionary standpoint. Recent studies conducted on understanding subsampling mechanisms in the visual and auditory domains (VanRullen et al., 2014; Zoefel et al., 2015) also support this idea. Unlike in vision, subsampling of auditory stimuli happens at the perceptual level and not at the source. Therefore, we can infer that high resolution temporal information of a sound is held in lower structures of the auditory processing pathway [discussed in introduction section 1.4A. Our results also demonstrate long term memory for these sounds, hinting at the existence of memory stores in lower areas of the auditory pathway.

III. Spatial correlates of memory for implicitly learned Gaussian sounds

3.1 Introduction

Results from the two previous studies described in my thesis highlight participants' ability to store short segments of acoustic information with high temporal resolution. These findings raise further intriguing questions regarding *how* and *where* this information is stored in the brain. In this experiment, several of these questions were tested using a combination of behavioral and imaging (fMRI) techniques.

First, we wanted to understand the inter-subject variability apparent during learning. It is evident from the results of Agus and colleagues fig 1.20A (Agus et al., 2010) and from our results (fig 2.5) that participants only learn a subset of the target cyclic noises – *old CNs* - that they are presented with. Previous studies on memory for Gaussian sounds have focused on testing memory for old sounds that were learned well, henceforth referred to as *best CNs*. However, the presence or absence of memory for old CNs that participants *did not learn* within the learning session, or *worst CNs,* has not been tested. Since learning is implicit, it is possible that participants learn acoustic features present in old CNs even when they demonstrate no improvement in discrimination performance during learning. These acoustic features may be subsequently consolidated during sleep [introduction section on sleep and memory]. On the other hand, since discrimination performance during learning (a') has been shown to significantly influence subsequent recall of intact and degraded versions of learned sounds (fig 2.8), it is also possible that features within old CNs that are not learned, are not stored. In this experiment, we investigated how implicit long-term recognition performance was different between best and worst CNs.

We also investigated the neural correlates of memory performance – i.e., areas of the brain involved in storing acoustic features present in old CNs. An interesting aspect of recognition memory for Gaussian sounds is the relationship between implicit and explicit processing. Participants are able to detect features in noise consciously (figure 1.20, figure 2.7), but do not seem to have conscious access to fine temporal resolution information and instead use vague descriptors for these features such as "whoosh", "clank" etc. The question of where these memories are stored is therefore fascinating. This, along with the importance of subcortical nuclei in auditory processing highlighted in introduction section 1.4 (describing in detail the auditory processing pathway and the role of each nucleus/area in processing information present in sounds), leads to several hypotheses regarding the role of these subcortical nuclei in memory for Gaussian sounds. However, due to technical difficulties, subcortical imaging in auditory processing has been rarely implemented. To our knowledge, the only fMRI study investigating neural correlates of memory for meaningless sounds using tone clouds (Kumar et al., 2014) focused on testing hypotheses regarding cortical correlates of memory, and demonstrated the role of the hippocampus and the auditory cortex in short term memory for meaningless stimuli. Another fMRI study (Griffiths et al., 2001) highlighted the importance of subcortical nuclei in temporal precision processing but did not test the role of these nuclei in storing acoustic features. If acoustic information is stored using a simple STDP mechanism, as suggested by the comparison of behavioral results with model predictions (section 2.5), high-resolution acoustic temporal information would be stored in subcortical regions. We were thus interested in

exploring the neural correlates of memory for Gaussian sounds, with an emphasis on understanding the role of subcortical nuclei.

Finally, we were also interested in testing – more exhaustively than previously – the influence of other parameters like musical expertise and sleep on implicit recognition memory performance.

Based on all findings described above, we tested the following hypotheses:

1) Long term, implicit memory for noise (demonstrated as a preferential bias to detect cyclic features in old vs. novel sounds) would be evident for old sounds that participants discriminate as cyclic with high probability during the learning session. However, acoustic features of old sounds that participants do not detect reliably during the learning session will not be stored, and participants will treat these sounds as novel sounds. This behavioral hypothesis is summarized in figure 3.1.
2) The ability to implicitly learn and recognize acoustic features at the individual level will depend on previous experience with music and on quantity of sleep between learning and testing sessions.
3) Both cortical and subcortical regions are involved in memory for meaningless stimuli. At the cortical level, hippocampus and regions of the auditory cortex will be involved in memory for old cyclic noises (Kumar et al., 2014). Modification in BOLD activation will be observed in response to learned acoustic information in sub-cortical regions, in accordance with predictions of STDP models of learning where feature selectivity occurs within a few layers of processing.

4) Specific hypotheses regarding each subcortical nucleus are based on detailed descriptions of the auditory pathway and plasticity observed in these nuclei (as described in the introduction section 1.4). These are listed below:

   a. Cochlear Nucleus (CN) – Based on the complexity of neuronal computations in this nucleus, we hypothesize that this nucleus would extract characteristic spectro-temporal features of cyclic sounds (discussed in intro section 1.4A). Neurons present in this nucleus have been shown to extract periodicity information from the incoming sound (Golding et al., 1995). Therefore, we predict that this nucleus would be differentially activated for cyclic and non-cyclic sounds.

   b. Superior Olivary Complex (SOC) – Neurons in this nucleus primarily compare inter-aural time and intensity difference of acoustic input. For binaural input, we do not expect to find any differences between old, novel and non-cyclic sounds.

   c. Inferior Colliculi (IC) – neural populations within sub divisions of the IC show diversity in response to the same stimulus. Previous research suggests that certain neurons in the IC store complex acoustic features and other neurons do not (Creutzfeldt et al., 1980). The global activation in response to old cyclic, novel cyclic and non-cyclic sounds is therefore difficult to predict with certainty, but it is likely that the IC is differentially activated for old cyclic and novel cyclic.

   d. Medial Geniculate Body (MGB) – Several electrophysiological studies on the MGB have consistently shown that neurons in this nucleus show plasticity and the ability to store memories. Feature extraction from the input stream occurs earlier in the pathway and the MGB receives individual spectro-temporal

features extracted by the cochlear nucleus. We therefore predicted that the MGB would be differentially activated for old cyclic and novel cyclic sounds but would not differentiate cyclic and non-cyclic sounds.

These hypotheses were tested using an implicit encoding and subsequent long-term implicit recognition paradigm, as previously described (Agus et al., 2010). To test the first hypothesis, in addition to the well learned old (*best* CNs) and novel sounds, participants were also presented with old sounds that they failed to learn (*worst* CNs), despite hearing both types of sounds the same number of times during the learning session. To test the second hypothesis, participants maintained a sleep diary during the intervening weeks between sessions and filled out a musical expertise questionnaire. To test the third and fourth hypotheses, participants performed the second session in an fMRI setting. Using a sparse acquisition fMRI paradigm during the implicit recognition test (Amaro et al., 2002) enabled participants to hear Gaussian noise segments within the scanner and discriminate sounds as cyclic or non-cyclic. Specific hypotheses regarding each nucleus in the auditory pathway were tested using region-of-interest based analysis, comparing fMRI blood oxygen level dependent (BOLD) activity for different kinds of trials.

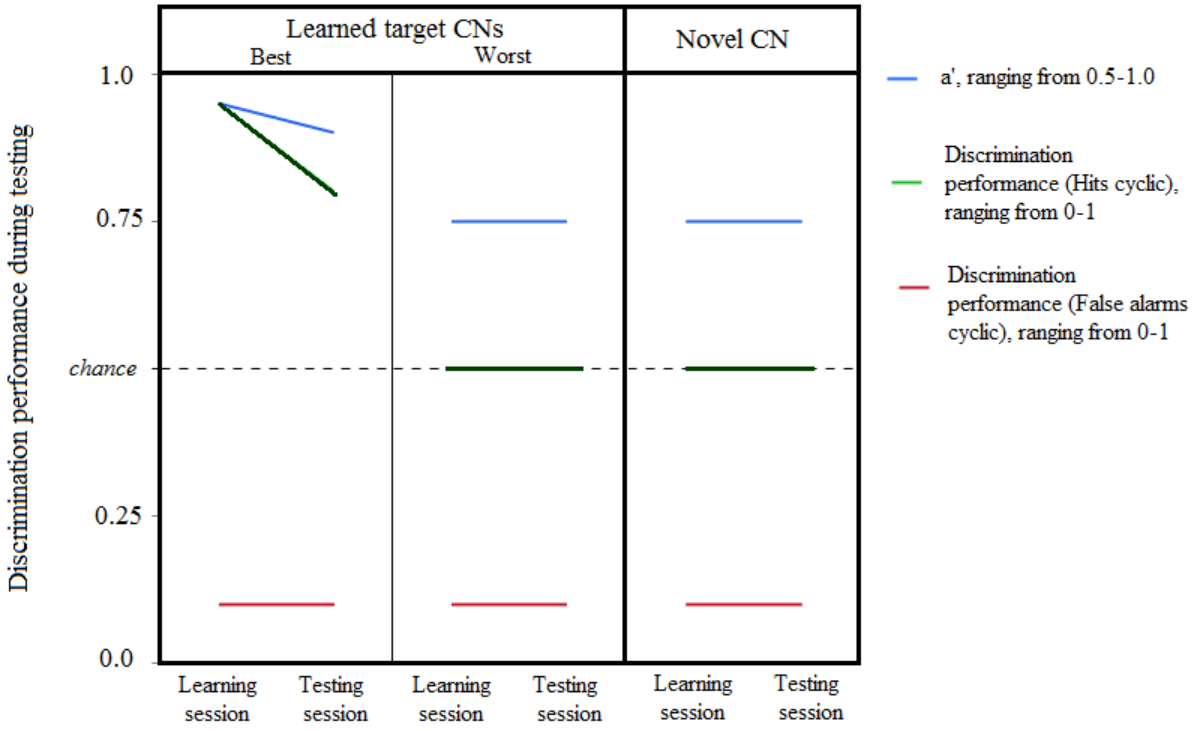


Figure 3.1 Predicted discrimination performance across the two sessions : in blue) a′ for old (best and worst) and novel CNs ranging from 0.5 to 1.0; in dark green) discrimination performance (hits) for old (best and worst) and novel CNs ranging from 0 to 1; in red) false alarms across the two sessions.

3.2 Materials and methods

3.2.1. Participants

A total of 21 healthy participants between 20 and 31 years of age were screened, of which 15 participants (mean age = 22.9, S.D = 2.6 years) went on to complete the whole experiment. All participants were compensated for their time with money; either a wire transfer of 100 euros for completing the whole experiment or with gift cards preloaded with monetary values of 10 euros (only screening) or 40 euros (screening and one experimental session). Participants were informed that the purpose of the experiment was to assess the neural correlates of auditory discrimination and were naïve to the actual hypotheses. All participants gave written informed consent in accordance to the declaration of Helsinki and the University of Toulouse and CNRS requirements for research with human participants [Protocol: CPP14-007a/2013-A01450-45].

3.2.2 Stimuli

Stimuli were programmed and generated using MATLAB R2013 (http://www.mathworks.com/). The sound stimuli were sequences of 16-bit pseudo-random numbers drawn from a normal distribution with a zero mean and played at a sampling frequency of 44.1 KHz. We constructed Cyclic noises (CN) and Non-Cyclic noises (N), both lasting 1 s in duration (audio samples can be found at http://m4.ups-tlse.fr/). A CN was generated as a 500-ms pseudo-random segment presented twice back to back (cycled) resulting in identical first and second halves of a 1-second

stimulus. An N was generated as a 1000-ms pseudo-random segment. Such Gaussian sounds have little variation of frequency over time resulting in flat spectrograms. Therefore, to illustrate these sounds, we plotted the actual amplitude variations over time for exemplars of CN (identical first and second halves) and N (Figure 2.1 A). All the stimuli were normalized to the hearing threshold of the participant (as determined from his/her audiogram, see the Procedure below) using the following formula:

$$X = x/rms(x) * (10^{((T+60)-94.6)/20)}) \qquad (2)$$

Where x is the segment of Gaussian sound, T is the hearing threshold (in dB) of either the left or the right ear, whichever is worse (equation 2), rms(x) is the root mean square of the segment and X is the normalized segment.

Throughout the experiment, participants were presented with 2 variations of CNs: a set of 10 uniquely generated CNs were presented several times to the participant during the learning and testing sessions (*old CNs)* while *novel CNs* were uniquely generated CNs that were only heard once throughout the experiment. All the Ns were uniquely generated and heard only once. The stimuli were presented to participants through headphones using either a psychoacoustic experiment programmed in MATLAB during the learning session, or the Presentation® software (Version 0.70, www.neurobs.com) during the fMRI testing session.

3.2.3 Task

Over two experimental sessions (learning and testing sessions), participants performed a forced-choice discrimination task. Each trial consisted in the presentation of a 1-second Gaussian noise (either CN or N) that the participant had to discriminate as cyclic/non-cyclic. Participants did not receive any feedback about performance during either session and were given breaks between blocks. All trials were presented in a randomized order. After session 1, each participant's performance was analyzed (as explained in the analysis section) and old CNs to be presented in the testing session were selected.

Additionally, participants also performed an explicit forced-choice recognition task at the end of the second session, where they had to indicate whether they had heard the sound before or whether it was new.

3.2.4 Procedure

Participants performed 2 experimental sessions approximately a month [mean = 31.8 days, S.D = 4.3] apart. MRI scanning was conducted during the second session only (testing session).

Session 1: learning

This session consisted of participants listening to sounds and performing the main forced-choice experimental task, while comfortably seated in front of a computer.

Screening - Before starting the learning experiment, all the recruited participants underwent a screening process. The hearing threshold for tones at different frequencies (0.25, 0.5, 1, 2, 4 and 8 KHz) was measured for both ears in an acoustic chamber (designed by studiobricks) using an audiometer (Materiel medical service SARL, France). The hearing threshold was calculated as follows.

$$T = \frac{t_{0.5}+t_1+t_2+t_4}{4} \quad (3)$$

Where t refers to tonal hearing thresholds (at 0.5, 1, 2 and 4 KHz respectively) and T is the average hearing threshold for the ear.

All the participants had normal hearing, with thresholds at/lower than 20 dB in both ears, and were included in the experiment.

Once included, all the stimuli for this participant were normalized using equations 1 and 2. Participants were then screened for eligibility to perform an MRI experiment. Finally, participants were screened for their ability to do the auditory discrimination task. They listened to samples of 5-s cyclic sounds constructed as 10 repeats of a 500-ms random noise segment and 5 second non-cyclic random noise sounds, until they could confidently differentiate between the two types of sounds. Participants who successfully completed the above screening stages went on to the training phase.

Training - Participants started the training phase by listening to a random ordering of 5 CNs (5s, 10 repeats of a 500-ms segment) and 5 Ns (5s of random noise). Following each sound presentation, participants had to indicate via a keyboard button

press if the sound was cyclic or not. After each trial, they were given feedback about their response. Once they had correctly identified all CNs, they moved to the next training stage, during which they were presented with 20 CNs (2s, 4 repeats of a 500-ms segment) and 20 Ns (2s). Once participants achieved a global accuracy of 80% of correct responses, they moved to the next training stage and were presented with 20 CNs (1.5s, 3 repeats of a 500-ms segment) and 20 Ns (1.5s) until they achieved a global accuracy of 70%. At any stage of the training, participants who did not reach criterion ended their participation in the study.

Implicit learning – After training, participants performed 10 blocks of the forced-choice discrimination task (as described earlier). In each block, participants were presented with 40 *Ns*, 20 *CNs* and 20 repeats of a unique *target CN*, in a random order. All participants heard the same set of 10 target CNs.

Session 2: Testing

This second experimental session was performed under MRI scanning conditions. The experiment included one run for auditory pathway localization and five runs during which the forced-choice discrimination task was performed. Each run lasted 12 minutes and included 80 trials. The localizer run included 40 noise sounds (Ns) trials and 40 silent trials. Participants were instructed to passively listen to the stimuli. In each forced-choice discrimination run, 40 CNs and 40 Ns were presented in random order. Over a total of 400 discrimination trials (5 runs), participants heard 4 old CNs 50 times each (which were selected from the 10 target CNs presented in the learning session), 100 novel CNs and 100 Ns. Old CNs were a subset of 4 CNs that participants heard during

the testing session chosen from the 10 target CNs participants heard during the implicit learning session. The 4 old CNs were selected for each individual as 2 *best* and 2 *worst* CNs, as explained below (Behavioral analysis section).

Immediately after MRI scanning, participants performed an explicit recognition memory task. They heard the 4 old CNs (2 best and 2 worst CNs) previously selected (old CNs) as well as 4 novel CNs, in a random order. Participants could listen to each of the stimuli as many times as they wished. For each sound, they had to indicate whether they had heard the CN during the course of the experiment or whether it was novel, and how confident they were in their response, ranging from 1 (completely unsure) to 5 (completely sure).

Questionnaires

To assess quality of sleep, participants were asked to maintain a sleep diary (Mary et al., 2013) during the 4 weeks between learning and testing sessions. For both experimental sessions, participants also filled out St. Mary's sleep questionnaire (Ellis et al., 1981) regarding their previous night's sleep quality . Lastly, to assess a possible influence of musical training on the ability to do the discrimination task, participants also filled out a slightly modified version of the Munich music questionnaire (Brockmeier et al., 2004), since this questionnaire was originally designed to assess musical ability following cochlear implantation.

3.2.5 Data acquisition and imaging protocol

Magnetic Resonance Imaging (MRI) was performed during the testing session using a 3 T clinical scanner (Philips Achieva, Best, The Netherlands) located in the hospital Purpan in Toulouse, France. High-resolution T1-weighted structural images were collected using a 3D sequence (multi-shot, in-plane resolution = 1 x 1 mm^2, slice thickness = 1 mm, TR = 8.1 ms, TE = 3.7 ms, flip angle = 8^o, TFE factor = 119, 170 transverse slices). Six functional runs, each including 81 T2*-weighted echo-planar scans (EPI), were completed using a sparse acquisition scheme (Amaro et al., 2002; Eden et al., 1999) (single-shot, TR = 9000 ms, TA (time of acquisition) = 2667 ms, TE = 30 ms, slice thickness = 4 mm, in-plane resolution = 2.4*2.4 mm^2, flip angle = 90^o, 36 transverse slices, SENSE factor = 1). Accordingly, each brain volume acquisition was followed by a silent period of approximately 6 s (with no acquisition). The 1-s stimuli were delivered during this period of silence. The timing of fMRI acquisition and stimulus presentation was optimized to nearly detect the peak of the hemodynamic response to the stimuli, while avoiding to detect the hemodynamic response to scanner noise (Amaro et al., 2002).

3.2.6 Analysis

Behavioral analysis:

Analysis of behavioral data from both sessions was performed using MATLAB and statistical tests were performed using JMP (Version *12.* SAS Institute Inc., Cary, NC, 1989-2007).

Following the learning session, the proportions of hits and false alarms in each block were calculated for each participant. The correct identification of a CN (novel and old CNs) as a cyclic noise was considered a hit and the incorrect identification of an N as a cyclic noise was considered a false alarm. In every participant, behavioral data were analyzed in each block by plotting hit rate for target and novel CNs versus false alarms rate (ROC graph). This graph allowed us to differentiate in each participant the target CNs that were the most and least accurately discriminated as cyclic noises. We individually selected 2 *best CNs* as those having the highest hits to false alarms ratio. We also selected 2 *worst CNs* as those having the lowest hits to false alarms ratio. These *best* and *worst* CNs were subsequently presented to the participant during the testing session.

Moreover, individual discrimination performance was computed for both experimental sessions using the principles of signal detection theory. In addition to hit rates, we calculated individual *a's*, a non-parametric measure of participants' sensitivity to differences between signal (target) and noise (distractor), i.e. cyclic vs. non-cyclic stimuli (Pollack and Norman, 1964; Stanislaw and Todorov, 1999). While sensitivity is traditionally evaluated using d', an assumption for using d' is that signal and noise distributions have equal standard deviations. In our experiment, since the signal trials include different subtypes of trials (CNs or old CNs) but not the noise trials (Ns), a' is a better estimate for sensitivity than d', as used in chapter 2 (equation 1, explained in chapter 2) (Viswanathan et al., 2016).

Data from the explicit task performed at the end of the testing session were analyzed by computing hit rates for previously heard old CNs and correct rejections of novel CNs. Using these values, we computed the participants' sensitivity to differences between signal (old CNs) and noise (novel CNs) as a' for explicit task, $a'_{explicit}$.

Neuroimaging analysis:

The MRI data acquired during the testing session were analyzed using SPM12 software (http://www.fil.ion.ucl.ac.uk, (Friston et al., 1995)). Two types of analysis were conducted. First, our specific *a priori* hypotheses regarding effects in the auditory pathway were tested using a region-of-interest (ROI) based analysis. We also performed a whole-brain exploratory analysis to investigate neural correlates of memory for noise in any brain regions.

Image pre-processing

All functional scans were corrected for physiological noise using the DRIFTER toolbox for SPM (Särkkä et al., 2012). This Bayesian method accurately tracks variations in cardiac and respiratory low frequencies using the principles of RETROICOR (Glover et al., 2000). Correction for low-frequency noise was applied since the BOLD signal from the brainstem is known to be particularly susceptible to physiological noise (Brooks et al., 2013). All functional scans were slice-timing corrected by phase shifting the component sine waves of the signal in a given slice, either forward or backward in time, up to the acquisition time of the reference slice (slice 16 of 36). Next, with the aim of

removing movement artefacts, the functional scans were realigned to the first volume using a 6-parameter rigid-body transformation calculated for each volume.

Further statistical analysis in *a priori* ROIs of subcortical regions (auditory brainstem and thalamic nuclei) were performed in subject's native space using these slice-timing, motion-corrected functional images. Thus, no further transformation or smoothing was performed for subcortical ROI analyses, as described recently (Aminoff and Tarr, 2015). For statistical analysis in *a priori* cortical ROIs (Heschl's gyrus and hippocampus), functional images were further processed exactly as in the whole-brain exploratory analysis (explained below).

For the whole-brain exploratory group analysis, the functional images were further co-registered to the anatomical T1-weighted image for each participant. Anatomical images were segmented based on inbuilt tissue probability maps for gray matter, white matter and CSF (SPM12) using affine registration and regularization. The segmented anatomical images were used to compute nonlinear transformations for normalization of the individual T1-weighted images to the MNI (Montreal Neurological Institute) T1 template. These transformations were applied to all functional images, which were resampled to 2 x 2 x 2 mm^3 voxels. Finally, functional images were spatially smoothed using an 8-mm full-width-at-half-maximum isotropic Gaussian kernel.

Statistical analysis

The General Linear Model approach (Friston et al., 1995) was used to perform individual analyses of functional data, as well as the whole-brain group analysis.

a. Individual analysis of the discrimination task

For each participant, an event-related model was designed, which included 4 regressors of interest corresponding to the 4 experimental conditions: best CNs, worst CNs, novel CNs and Ns. In each regressor, onsets of stimuli were modeled as delta functions of zero duration convolved with the canonical hemodynamic response function (hrf). Participant's motion was modeled using 6 additional confounding regressors. A high-pass filter (cut-off = 128 s) was applied to remove slow signal drifts.

b. Subcortical ROI analysis (group)

Functional ROIs evidencing the subcortical auditory pathway were created for each individual. To this aim, stimuli onsets for the two conditions (Ns, silence) used in the localizer functional run were modeled using delta functions convolved with the canonical hrf. Six movement parameters (calculated for each trial during pre-processing) were added as confounding regressors to partially account for variance in the signal due to subject's motion. The data were also high-pass filtered with a frequency cut off of 128 s. An explicit mask created from the structural scan of the participant was used to limit the analysis to participant's brain. The model was estimated on the DRIFTER corrected, slice-timing corrected, realigned functional images from the localizer run and SPMs contrasting 'Ns' and 'silence' trials were

created. Using clusters of activity in each subject's native space at a threshold of $p < 0.01$ (uncorrected) and based on anatomical landmarks, we identified bilateral brainstem and thalamic nuclei of interest, i.e. cochlear nuclei (CoN), inferior colliculi (IC) and medial geniculate bodies (MGB). The use of a low threshold to detect subcortical nuclei activity was justified due to the small size and low BOLD response in the nuclei (Moerel et al., 2015). Note that all nuclei could not be localized in every participant (see Results section). Moreover, clusters of activity were hardly evidenced in the region of the superior olivary complex and lateral lemniscus (these 2 nuclei could not be distinguished on functional maps). Therefore, further analysis could not be conducted on these *a priori* ROIs. Using the MARSBAR toolbox (Brett et al., 2002) in SPM8, 3D spherical ROIs centered on individual local peaks of activity were created for each participant. The radii and structural landmarks for identifying functional activations for different brainstem and thalamic ROIs were determined based on stereotaxic atlas (Olszewski and Baxter, 2014) as well as anatomic and physiological research on these nuclei (Moore, 1987; Winer, 1984), i.e. 3 mm for the CoN (1.5 mm each for dorsal and ventral CoN) located at the base of the pons (which interfaces with the medulla oblongata), 3.5 mm for the IC located within the clearly visible corpora quadrigemina located on the tectum, and 3 mm for the MGB which are dorso-lateral to the tectum and close to the brachium (white matter). [These regions are highlighted in the introduction section on the ascending auditory pathway 1.4A, figure 1.13]. Using MarsBar, individual models for the discrimination task were estimated on ROI voxels, and mean parameter estimates for contrasts of interest were computed for each participant and each ROI.

c. Cortical ROI analysis

–Analyses in Heschl's gyrus (HG) and hippocampus (HC) ROIs were conducted in standard space. Regarding the HG region, it was difficult to locate based on analysis of individual functional localizer scans, as large clusters of activity (encompassing both primary and secondary auditory cortices) were found in the superior temporal area. Thus, the model was estimated on *structural* ROIs with the DRIFTER corrected, slice-timing corrected, realigned, normalized and smoothed functional images. Bilateral HG and bilateral HC ROIs were selected from the MarsBar's AAL ROI library in standard MNI space (Tzourio-Mazoyer et al., 2002). Individual models for the discrimination task were estimated on ROI voxels using MarsBar and mean parameter estimates for contrasts of interest were computed for each participant and each ROI.

d. Whole-brain exploratory analysis

For this exploratory analysis, one-sample t-tests were modeled for contrasts of interest, using a random-effects procedure. Individual participants' performances (hit rates) in the discrimination task were included as a confounding covariate.

For example, to analyze the whole brain activation differences seen for Old CN compared to Novel CN, second level analysis was performed on all the participants for this contrast. The behavioral difference for each participant (the mean hits for old CNs – mean hits for novel CNs) was included in the model as a confounding covariate. This allowed us to perform weighted averaging at the group level, where data from participants with a larger behavioral difference in the two conditions were weighted more than participants with a smaller behavioral difference between the two conditions.

e. Selection of contrasts

Based on the four trial types (best CNs, worst CNs, novel CNs and Ns), five main contrasts of interest were defined:

(i) **(Old – New)** was defined to visualize the areas involved in memory for Gaussian noise. Best and worst CN trials were combined and contrasted with novel CN and N trials. Note that this contrast may be sensitive to sound cyclicity effects.

(ii) **(Old CNs – Novel CNs)** was defined to visualize the areas involved in storing features of cyclic sounds as a function of repeated exposure, irrespective of behavioral response. Best and worst CN trials were combined and contrasted with novel CN trials. This contrast was aimed at elucidating regions implicated in storing idiosyncratic acoustic features in old CNs.

(iii) **(CNs – Ns)** was defined to visualize the areas involved in detecting noise cyclicity. Best, worst and novel CN trials were combined and contrasted with N trials. Note that this contrast may also be sensitive to memory effects.

(iv) **(Novel CNs – Ns)** was defined to visualize areas involved in detecting cyclicity in novel sounds.

(v) **(Best CNs – Worst CNs)** was defined to visualize areas involved in storing specific acoustic features that are preferentially stored. Best CN trials were contrasted with worst CN trials.

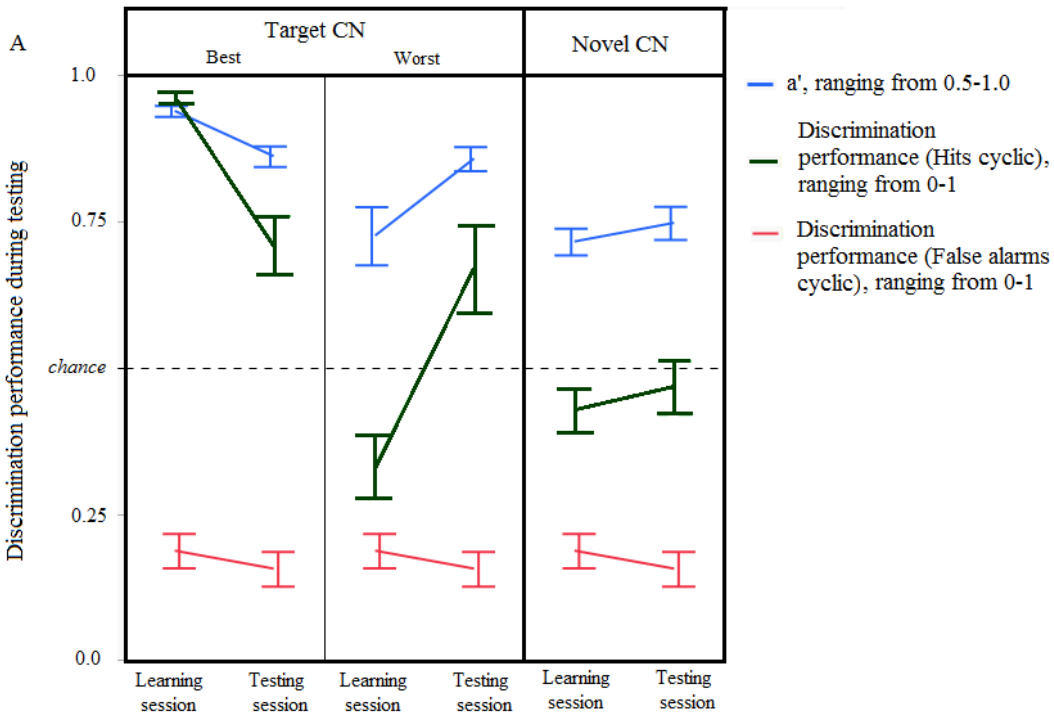
3.3 Results

3.3.1 Behavioral results

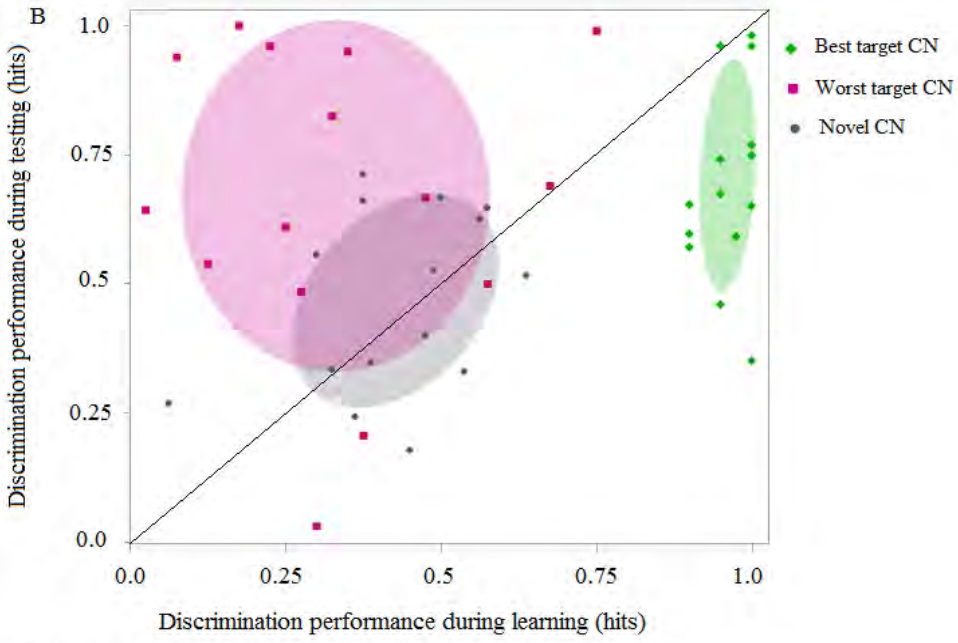
Individual a' values for the 15 participants who completed the whole experiment [inclusion criteria: a' > 0.5 during the learning session] were equivalent across the two sessions [$F(1,90) = 1.09$, $p = 0.299$], indicating that participants' discrimination strategy was comparable across sessions, demonstrating that participants do not improve at the task in general. Correct discrimination of different types of cyclic sounds (proportion of hits) was therefore compared across sessions using a two-way repeated-measures ANOVA, testing main effects and interaction of within-subjects factors 'session' (2 levels, 'learning' and 'testing') and 'CN type' (3 levels, 'best CN', 'worst CN' and 'novel CN'). Hits for CNs remained equivalent across sessions ($p = 0.257$) but CN type was a significant predictor of performance [$F(2,90) = 44.32$, $p < 0.0001$]. Tukey's Honestly Significant Difference (Tukey's HSD) post-hoc test showed that discrimination of best CNs were higher than discrimination of novel CNs [effect size ($mean_{(i)}$-$mean_{(j)}$) = 0.388, $CI_{95\%}$ = (0.28, 0.49), $p < 0.0001$], and discrimination of worst CNs [effect size = 0.34, $CI_{95\%}$ = (0.23 , 0.44), $p < 0.0001$]. Discrimination of Novel CNs and Worst CNs was equivalent [$p = 0.465$]. Interestingly, there was a session × CN type interaction, that is, the discrimination performance for different types of CN depended on the session [$F(2,90) = 21.59$, $p < 0.0001$]. Tukey's HSD showed that discrimination of novel CNs were equivalent across sessions [$p = 0.9891$], indicating that participants did not show improvement in the task. Discrimination performance for best CNs was lower during

the testing session [effect size = -0.251, $CI_{95\%}$ = (-0.44, -0.07), p = 0.0023], while discrimination of worst CNs was higher during the testing session [effect size = 0.337, $CI_{95\%}$ = (0.15, 0.52), p < 0.0001] when compared to the learning session. Lastly, to eliminate the possibility that the observed preference to detect cyclicity in target CNs during the testing session was due to within-session learning, discrimination performance for the first 10 target and 10 novel CN trials were compared. From the first trial of the first block during the testing session, participants discriminated old (best and worst) target CNs better than novel CNs [repeated measures ANOVA comparing mean discrimination rate over first 10 trials of the testing session: $F(1,406) = 18.87$, $p < 0.0001$]. These behavioral results are summarized in figure 3.2.

Additionally, similar to our previous behavioral experiment (chapter 2), analysis of training data using a two-way repeated measures ANOVA showed no effect of number of training stages (just before the learning session) on discrimination performance during testing [stage 1- $F(3,15) = 3.6756$, $p = 0.1204$; stage 2 – $F(2,15) = 2.6743$, $p = 0.1831$ and stage 3 – $F(2,15) = 1.4425$, $p = 0.3375$].



Each error bar is constructed using 1 standard error from the mean.



Density ellipse (50% coverage)

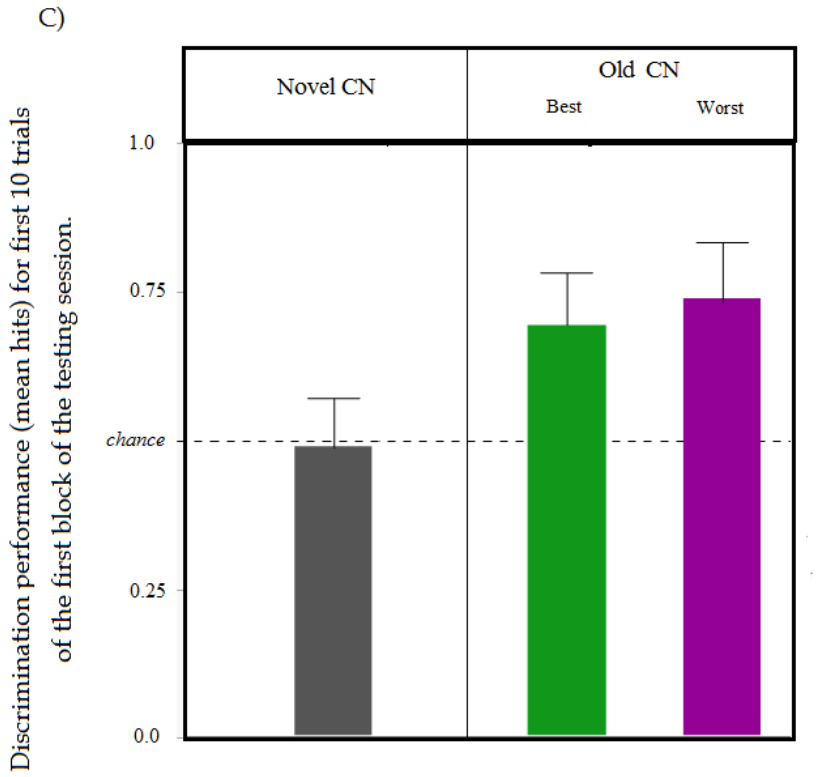


Figure 3.2 A) Discrimination performance across the two sessions (n=15): in blue, a' for old (best and worst) and novel CNs; in dark green, discrimination performance (proportion of hits) for old (best and worst) and novel CNs; in red, proportion of false alarms across the two sessions. B) Relationship between discrimination rates for different types of CNs - old (best and worst) and novel CNs, across the two sessions. Points above the diagonal indicate participants with higher discrimination rates during the testing session for each trial type. Note that discrimination of Novel CNs is close to the diagonal (equivalent performance across sessions) for most participants. C) Discrimination performance (hits) during the first 10 trials of the testing session for old (best and worst) as well as novel CNs for all participants. From the first trial of the first testing block, participants were better at discriminating old compared to novel CNs.

Next, behavioral data from the explicit task were analyzed. Correct response rates for old CNs (mean = 0.49, SE = 0.025, $CI_{95\%}$ = (0.448, 0.549)) and for Novel CNs (mean = 0.49, SE = 0.036, $CI_{95\%}$ = (0.422, 0.575)) were at chance level. The self-reported confidence in response could have affected the response of participants (i.e., accuracy should be proportional to confidence). Before running an ANCOVA (analysis of covariance), we confirmed the absence of interaction between trial type ('Old CNs' and 'Novel CNs')

and confidence [$F(1,45) = 0.5096$, $p = 0.479$]. Two way repeated measures ANCOVA testing the main effect of trial type [$F(1,45) = 0$, $p = 0.997$] and confidence [$F(1,45) = 0.005$, $p = 0.942$] showed an absence of effect of either variable on the performance. At the individual level, no participant was able to reliably distinguish old from novel CNs. The source of variability in the data was between participants with some responding 'old' for all CNs and others responding 'novel' for all CNs.

To better understand the improvement in discrimination performance for worst CNs during the testing session, a regression to the mean analysis was run. Regression to the mean (RTM) (Barnett et al., 2005) is a statistical phenomenon which emerges when repeated measurements are made on the same variable. Assuming that the variable has a normal distribution, successive observations are likely to move closer to the 'true' mean of the distribution. If the first observation is close to the true mean of the distribution, following observations do not result in noticeably different values. However, if the first observation is away from the true mean, following observations have a higher likelihood to shift towards the true mean.

Results of the RTM analysis applied to our data are shown in figure 3.3. In each participant, difference in cyclicity detection (hit rate) was calculated for all old and novel CNs, across sessions. Here, the hit rate in the learning session represents the 'first observation' for a given subject; the hit rate in the testing session is 'following repeated observation'. Differences in discrimination performance demonstrated an RTM effect for worst CNs [$R^2 = 0.3427$, slope = -1.031, intercept = 0.6787, $p = 0.0007$], but not for best CNs [$R^2 = 0.0243$, slope = -0.876, intercept = 0.5875, $p = 0.41$] or novel CNs [R^2 =

0.2269, slope = -0.572, intercept = 0.2885, p = 0.0727]. These results imply that the 'poor' performance observed for worst CNs during the learning session reflects a measurement error, and that acoustic features in both best and worst CNs are learned as a consequence of number of presentations.

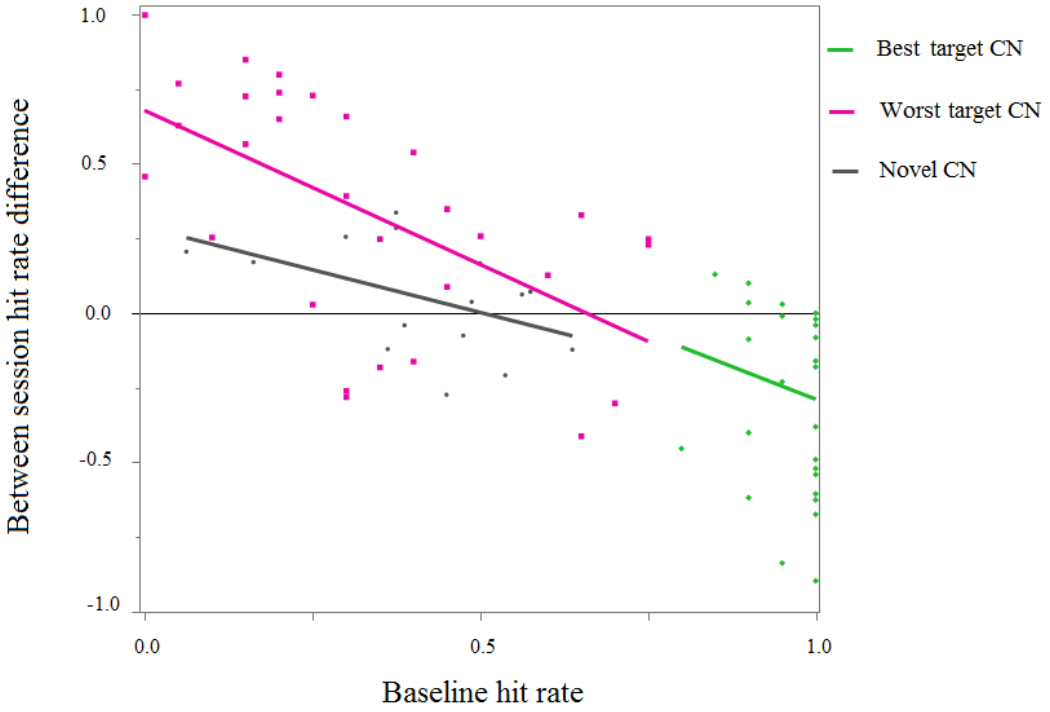


Figure 3.3 Regression to the mean analysis applied to behavioral results demonstrating a RTM effect in cyclicity detection across sessions for worst CNs, but not best CNs or novel CNs. 'Baseline hit rate' is the hit rate during the first observation which is the learning session.

3.3.2 Results of self-report sleep and musical questionnaires

Participants were asked to fill out three sets of questionnaires, the results of which are summarized in table 4.1. While several parameters were assessed, very few effects were actually observed. Musical training did not correlate with discrimination performance,

but frequency of singing did correlate with discrimination performance during the testing session. Sleep latency – or the time to fall asleep - the night before testing session also correlated with discrimination performance during the testing session. Although we intended to collect values for many variables that could have impacted sound discrimination and retention in memory, these few correlations are difficult to interpret. Replication on a larger sample size or with use of more objective measures might be necessary before the influence of sleep, musical training, caffeine and nicotine on task performance are understood.

Category	Question	Levels	Session	F ratio	P value
Effect of musical training on a'	Duration of musical training	0 , less than 3 years, greater than 3 years	Learning	1.688	0.226
			Testing	0.7549	0.4911
	Hours per day participants spent listening to music	0.5-1 hour, 1-2 hours, greater than 2 hours and all day	Learning	2.1105	0.1569
			Testing	0.6358	0.6073
	Frequency of playing an instrument	very high, high, moderate, low	Learning	0.4741	0.7066
			Testing	0.5437	0.6602
	self-rated importance of music	very high, high, moderate	Learning	1.175	0.342
			Testing	0.648	0.5405
	frequency of singing	very often, often, sometimes, rarely	Learning	2.1858	0.1528
			Testing	4.6404	0.031*
Effect of sleep before sessions on a'	Sleep quantity (hours) the day before learning session	Continuous variable, min reported value 4, max 9	Learning	3.3055	0.0922
			Testing	3.0292	0.1054
	Sleep quantity (hours) the day before testing session	Continuous variable, min reported value 4, max 9.	Testing	0.4811	0.5001
	Sleep quality the night before learning session	Discrete values ranging from 1 to 8	Learning	0.0599	0.8104
			Testing	0.0053	0.9432
	Sleep quality the night before testing session	Discrete values ranging from 1 to 8	Testing	0.7722	0.3955
	Sleep latency the night before learning session	Continuous variable, min reported value 5, max 40	Learning	0.0041	0.9498
			Testing	0.0038	0.9519
	Sleep latency the night before testing session	Continuous variable, min reported value 5, max 45 minutes	Testing	6.1805	0.0273*
	Average sleep per day during the intervening 4 weeks	Continuous variable, min reported value 6, max 9 hours per day	Testing	0.0409	0.8428
Effect of other parameters on a'	Average nicotine intake per day during the intervening 4 weeks	Continuous variable, min reported value 0, max 6.5 cigarettes/day	Testing	0.2443	0.6294
	Average caffeine intake per day during the intervening 4 weeks	Continuous variable, min reported value 0, max 4 cups/day	Testing	0.4007	0.5377

Table 3.1 : This table summarizes the effects of sleep, musical training, nicotine and caffeine on discrimination performance during the learning and testing sessions.

3.3.3 ROI results

As previously described in the Methods, individual functional brainstem and thalamic ROIs were created from the analysis of the localizer scan. At a threshold of $p < 0.01$, not all a priori subcortical areas were localized in all participants. Left CoN was localized in 9, left IC in 10 and left MGB in 11 (of 15) participants. Right CoN was localized in 9, right IC in 9 and right MGB in 10 (of 15) participants. Therefore, group effects reported below in each of these ROIs were computed for the number of participants where the ROI could be identified. Regarding cortical ROIs (based on the AAL atlas), group effects were computed on all 15 subjects.

One-sample t-tests were run for the percentage BOLD signal change for each of the contrasts in each ROI against a theoretical mean of 0 (null hypothesis that the signal does not change in the ROI between the 2 conditions that are contrasted). These results along with ROIs in a representative subject are reported below for each contrast and are highlighted in figure 3.4 (subcortical ROIs, native space) and figure 3.5 (cortical ROIs, normalized MNI space).

(i) Old – New (memory effects, although sensitive to sound cyclicity)

At the subcortical level, the right CoN showed higher activity in response to New vs. Old sounds [mean = 9.881, SE = 3.007 , $t(9) = 3.2856$, $p = 0.0111$] whereas the left CoN showed higher activity in response to Old vs. New sounds [mean = -6.421, SE = 2.728, $t(9) = -2.3533$, $p = 0.0464$]. Additionally, the right inferior colliculus [mean = 7.046, SE = 2.906, $t(9) = 2.4246$, $p = 0.0415$] showed higher activity in response to Old vs. New

sounds. At the cortical level, activity in the left hippocampus [mean = 1.36, SE = 0.584, t(15) = 2.3195, p = 0.0353] was increased in response to Old vs. New sounds.

(ii) Old CNs – Novel CNs (memory effects for specific features in old CNs)

At the subcortical level, two regions showed higher activation in response to old vs. novel CNs, namely the left cochlear nucleus [mean = 9.993, SE = 3.421, t(9) = 2.9209, p = 0.0193] and the right medial geniculate body [mean = 5.65, SE = 2.22, t(10) = 2.545, p = 0.0315]. At the cortical level, both the left [mean = 1.585, SE = 0.688, t(15) = 2.3029, p = 0.0371] and the right [mean = 1.087, SE = 0.507, t(15) = 2.1449, p = 0.05] hippocampi evidenced a similar pattern of activity.

(iii) All CNs – Ns (detection of sound cyclicity, although sensitive to novelty)

From all subcortical and cortical a priori ROIs, only the left cochlear nucleus showed differential BOLD response for cyclic and non-cyclic sounds. This BOLD response was increased for cyclic sounds [mean = 9.656, SE = 3.654, t(9) = 2.6423, p = 0.0296].

(iv) Novel CNs – Ns (detection of sound cyclicity in novel sounds)

No differential activations were found in any of the subcortical [left and right CoN, IC, MGB; p >= 0.05] or cortical [left and right HC and HG; p >= 0.05] sites.

(v) Best CNs – Worst CNs (detection of acoustic features that are learned)

Despite no behavioral differences between the two conditions, to test a priori hypotheses we looked at differences in activation between best and worst CNs. No

differential activations were found in any of the subcortical [left and right CoN, IC, MGB; $p >= 0.05$] or cortical [left and right HC and HG; $p >= 0.05$] ROIs of interest.

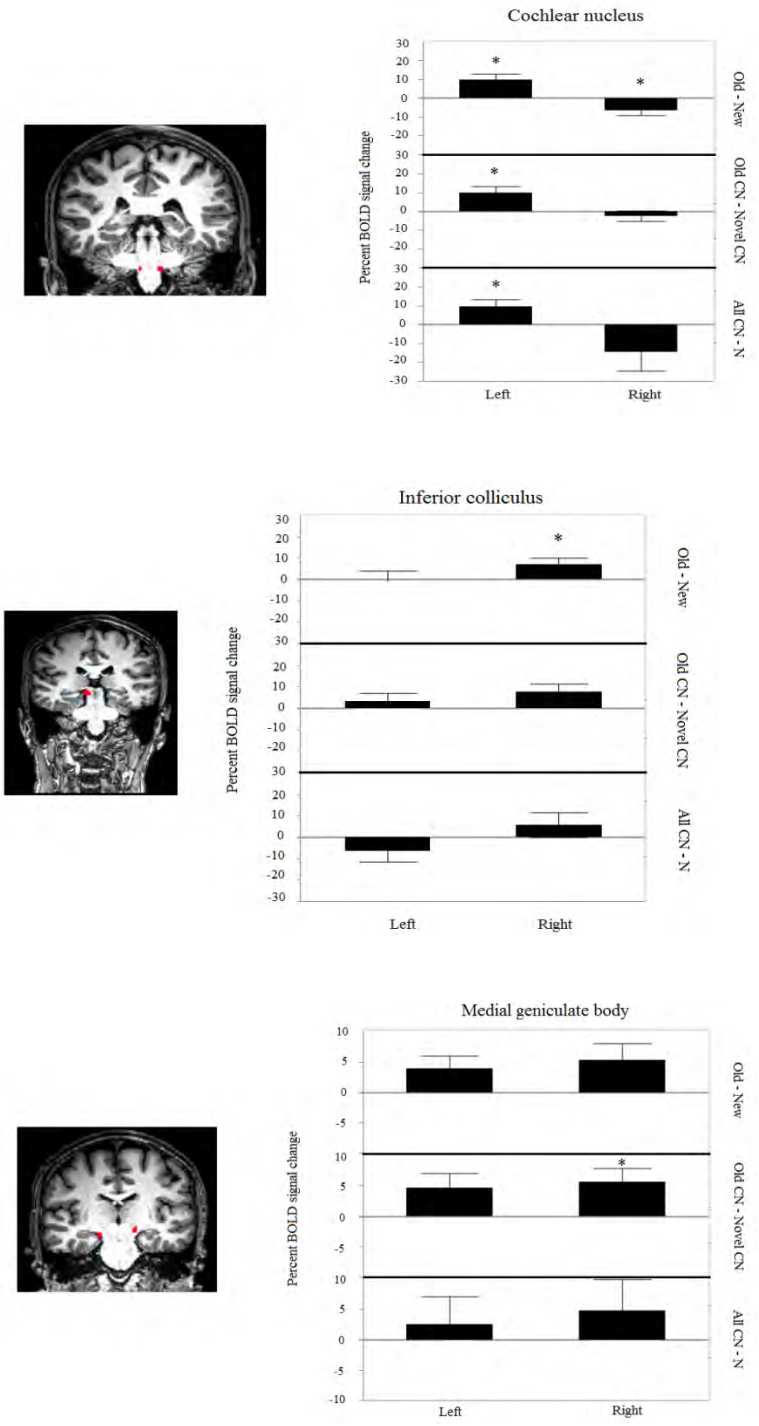


Figure 3.4 Percent BOLD signal change in subcortical ROIs. Left: examples of functional ROIs constructed using MarsBAR in a representative subject. From top to bottom – (i) left cochlear nucleus ROI identified from functional data and overlaid on the participant's anatomical scan, (ii) left inferior colliculus ROI identified from functional data, (iii) left medial geniculate body ROI identified from functional data. Right: corresponding BOLD signal change observed at the group level for the 3 contrasts of interest.

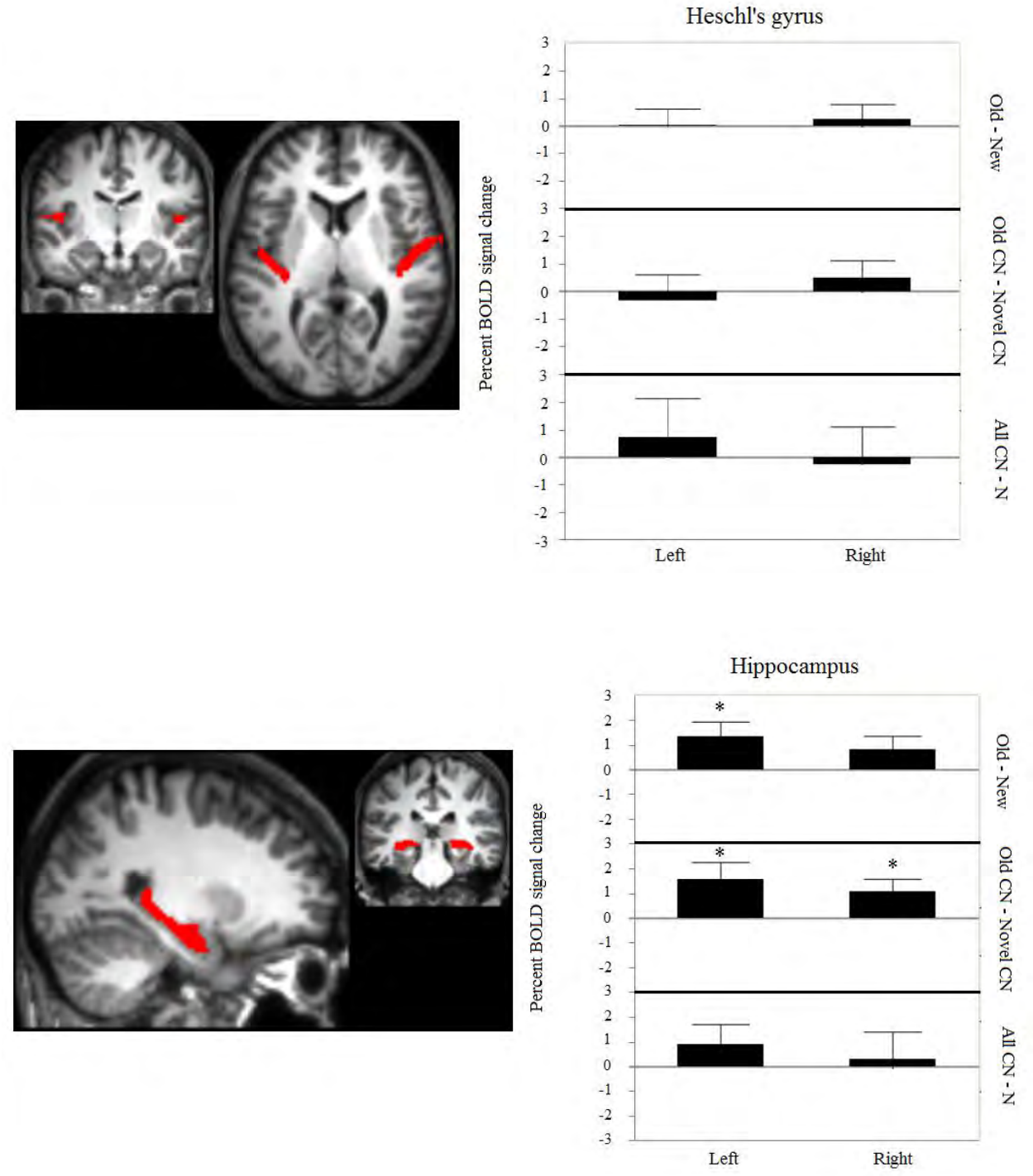


Figure 3.5 Percent BOLD signal change in anatomical cortical ROIs. Left: structural ROIs from AAL atlas in MNI space. From top to bottom – (i) bilateral Heschl's gyrus (ii) bilateral hippocampus. Right: corresponding BOLD signal change at the group level for the 3 contrasts of interest.

3.3.4 Summary of the ROI results (by region)

The cochlear nucleus seems to be involved in both memory functions and detection of cyclicity. The inferior colliculus and medial geniculate are involved in memory but not in detection of cyclicity. Interestingly, the medial geniculate body might be actually storing features of learned cyclic stimuli, resulting in the differentiation of (highly perceptually similar) old and novel CNs. A lateralization of information transfer was observed, with bilateral activations at the level of the cochlear nucleus and unilateral activations at all higher ROIs in the ascending pathway. At the cortical level, the hippocampus was selectively activated only for memory contrasts. However, Heschl's gyri (left and right) were not differentially activated for any of the contrasts. Corresponding to behavioral results, no activation differences were observed between best and worst CNs in any of the ROIs of interest.

3.3.5 Whole brain results

Results from the whole brain exploratory analysis are listed in table 4.2 and illustrated in figure 3.6.

Contrast	Brain areas	# of voxels	Peak MNI coordinates			T value	p value
			x	y	Z		
Old - New	posterior cingulate	31	8	-36	38	5.553	< 0.001
	gyrus cerebellum	36	20	-76	-28	6.67	< 0.001
Old CN - Novel CN	medial geniculate body	17	-6	-26	-14	5.2	< 0.001
	posterior thalamus	77	-16	-28	14	6.49	< 0.001
	cerebellum	88	18	-66	-22	8.03	< 0.001
	posterior cingulate gyrus	48	16	-36	44	6.39	< 0.001
	central gyrus	30	-4	-34	48	6.36	< 0.001
	pre-central gyrus	84	-36	-56	60	8.12	< 0.001
	inferior frontal gyrus	10	-26	48	20	6.54	< 0.001

Table 3.2: List of all clusters of activation for all contrasts of interest (cluster size > 10, p < 0.001 uncorrected).

In agreement with the results obtained from the ROI analysis, the left medial geniculate body was more activated in response to Old CNs vs. Novel CNs. Additionally, the posterior cingulate cortex, which has long been identified to have connections with the hippocampus and thalamic regions as part of the Papez circuit (Papez, 1937), was also activated in response to old compared to novel stimuli. Areas of the cerebellum were activated for both contrasts investigating memory effects (Old-New and Old CN – Novel CN). No supra-threshold activations were observed for contrasts investigating effects of sound cyclicity in the whole brain analysis. Similarly, no supra-threshold activations were observed for contrasts investigating differential encoding of features present in best and worst CNs.

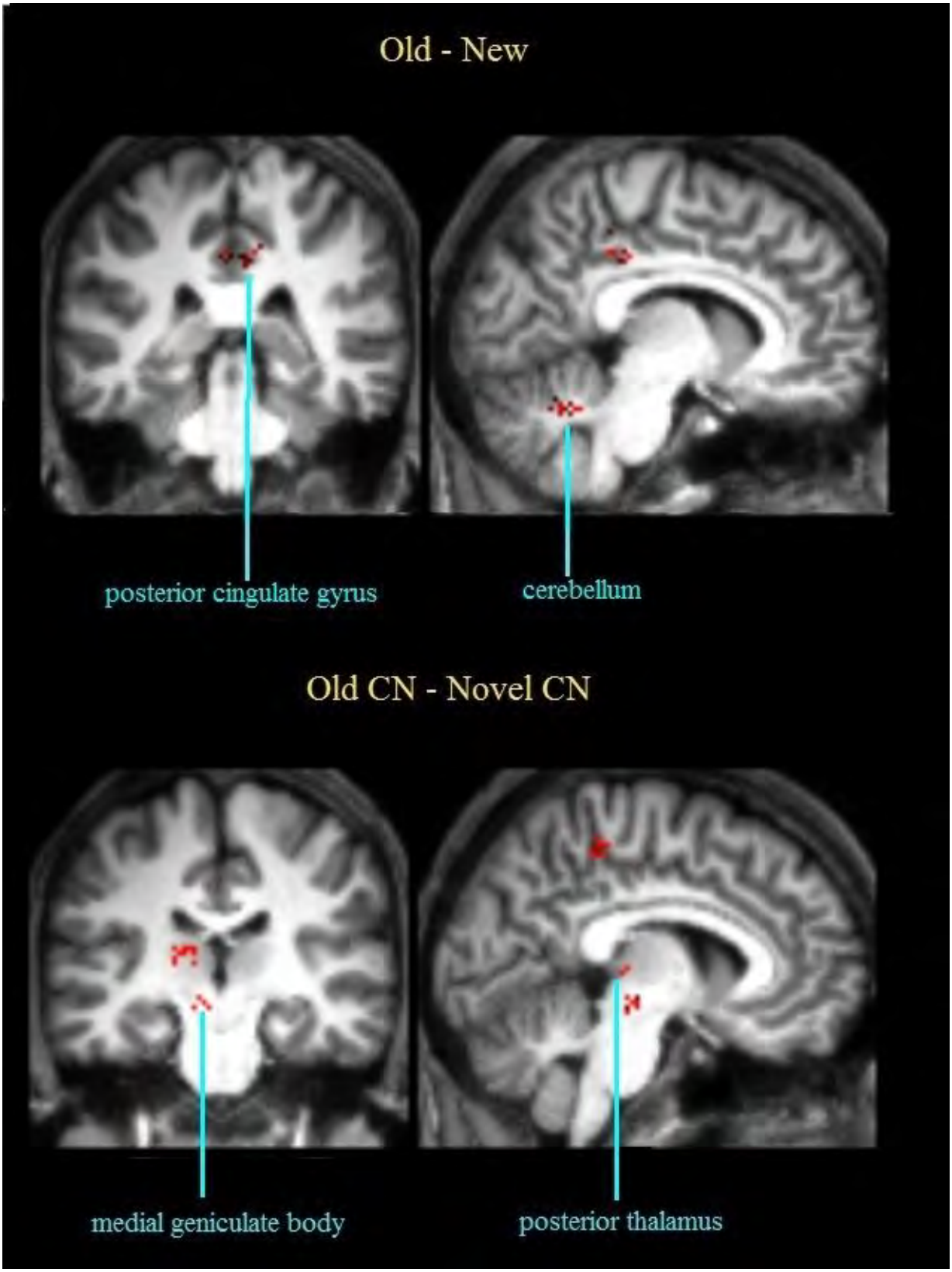


Figure 3.6: This figure shows group activations for the contrasts of interest (T = 3.85, cluster size > 10, $p < 0.001$ uncorrected). Group activity is overlaid on one participant anatomical scan, normalized into standard MNI space.

3.4 Discussion

3.4.1 Significance of behavioral results

The behavioral results of this experiment demonstrate that memory for features present in Gaussian sounds can be learned and retained over several weeks. While the discrimination performance was as predicted for best CNs (refer figure 3.1), surprisingly, participants also demonstrated memory for worst old CNs. Since detection rate for novel CNs, as well as rates of false alarms, were equivalent across sessions, participants' performance was not improving overall, which clearly argues in favor of true memory for both best and worst CNs. In fact participants' ability to correctly discriminate a worst CN as cyclic seemingly 'improves' across sessions and is well above chance level in the testing session (fig 3.2.A). On closer investigation however, regression to the mean analysis showed that this apparent improvement in detection of worst CNs across sessions is a statistical phenomenon. That is, the 'poor' learning observed for worst CNs (by figure 1.20A (Agus et al., 2010) and fig 2.5 demonstrating variability in learning) is not a real index of true encoding and probably just reflects random measurement error during the learning session. This finding implies that acoustic features in old CNs are encoded as a function of repetition irrespective of conscious detection of cyclicity. This is consistent with the absence of fMRI activity differences in response to best CNs and worst CNs, as observed in our data. Results from chapter 2 (fig 2.8) also demonstrated that inter-individual differences

in discrimination sensitivity during learning affected subsequent recall. As the first study to test memory for 'poorly' learned CNs, these results demonstrate purely statistical learning in the auditory modality, similar to unsupervised learning in the STDP model (Masquelier et al., 2008, 2009, 2016). Putting these findings together, it appears that implicit recognition memory for Gaussian noise depends on participant's sensitivity to differences between acoustic features, but not on measures of *conscious* behavioral detection of these features. This lack of conscious access to stored acoustic features is confirmed by participants' lack of explicit memory for old CNs, as observed here.

Exhaustive measurements (self-report) of participants' sleep quantity and quality between the learning and testing sessions only showed a weak link between sleep on the day before testing and discrimination sensitivity during the testing session. No other links between sleep and discrimination sensitivity post-learning were apparent from our data. These findings contradict what is already known about the influence of sleep and prior training on memory (introduction section 1.2B). More stringent measures, for example sleep assessed using brain electrical activity using electroencephalography or muscle activity using actiwatch, a wearable device used to accurately quantify sleep in terms of muscular activity during sleep and daytime activity, (http://www.actigraphy.com/) are necessary to understand the true relationship between sleep and memory for Gaussian sounds. While less comprehensive than EEG measures of sleep, actigraphy is a more practical way of accurately quantifying some aspects of sleep in long term memory studies.

Data from a questionnaire assessing musical expertise using a standard auditory assessment test did not demonstrate any correlation between discrimination sensitivity (a') and years of musical training. Meanwhile, frequency of singing did weakly influence discrimination sensitivity in the testing phase. However, as highlighted by some researchers, self-report measurements do not take into account participants with musical ability but no formal training, as well as people with extensive musical training who may not be very skilled (Law et al., 2012). More quantitative measurements of musical ability might reveal the influence of this factor on implicit recognition memory for acoustic features. One such battery creates a profile of music perception skills (PROMS) for each participant to measure music perception across multiple domains such as tonal, temporal, dynamic and qualitative (Law et al., 2012). Such tests might help to elucidate the precise relationships between proficiency in music/sound processing and memory for Gaussian noise.

3.4.2 Significance of ROI results

Despite several hypotheses regarding the role of subcortical nuclei in auditory processing, this is the first study to systematically test the role of these nuclei in storing meaningless feature information in long-term memory. While results from chapter 2 alluded to the possibility that features can be stored at the subcortical level, the results of ROI analysis confirm this. First, the CoN was found to differentiate cyclic and non-cyclic sounds as predicted (hypothesis 4.a). Interestingly, we did not find the CoN differentially activated for novel CNs and Ns. This might be due to a lack of response

on 'average' in the CoN in response to novel stimuli; in other words, perhaps the activations invoked by novel CNs and Ns were variable across trials resulting in no global differences between these two. Surprisingly, the CoN could also differentiate old CNs from novel CNs and Ns, implicating it in memory. To our knowledge, memory has never been investigated in the CoN. Since certain types of neurons present in this nucleus extract spectro-temporal features present in incoming sounds, detection of acoustic features was expected, but memory for these acoustic features over features present in novel CNs and Ns is surprising. However, the corticofugal descending fibers have been shown to modulate plasticity at the subcortical level (introduction section 1.4A). Therefore while this finding is surprising, the existence of memory in the CoN is computationally feasible. However, since the CoN was close to the edge of the field of view of the functional scans, the role (or lack thereof) of the CoN in detecting cyclicity and/or storing acoustic features is harder to definitively interpret from this dataset. Second, the right IC was involved in differentiating old from novel sounds. It was difficult to make precise hypothesis regarding the role of IC in this task (hypothesis 4.b). Thus, the next node in the auditory processing pathway did not differentiate cyclic from non-cyclic sounds but did recognize old sounds as being different from novel ones. The last step in auditory subcortical processing occurred in the ipsilateral MGB, hypothesized to store acoustic information in long-term memory (hypothesis 4.c). Indeed, neurons in the right MGB were able to differentiate highly perceptually similar old and novel CNs as predicted.

Regarding cortical ROIs, we observed that, as anticipated (hypothesis 3) and found previously (Kumar et al., 2014), the hippocampus may be involved in memory for

Gaussian sounds, as suggested by activity differences in response to old CNs ,novel CNs and Ns. Further, the hippocampus was not involved in the detection of cyclicity. Only a couple of studies have implicated the hippocampus in detecting 'low level' changes in perceptual features in vision (Strange et al., 1999) and audition (Kumar et al., 2014). In the Strange et al. study, authors demonstrated the involvement of hippocampus in differentiating novel fonts. In the Kumar et al. study the hippocampus was shown to be involved in short-term memory for features in meaningless tone clouds, highly consistent with our results. Kumar et al. proposed that the hippocampus would convert representations in the primary auditory cortex into sparser forms. However, an alternate explanation of hippocampal function is possible in agreement with all these findings: if the hippocampus was recruited to form a memory trace based on functional requirements of the task at hand (rather than merely increasing complexity of the stimuli), it would be subsequently involved in implicit and/or explicit recognition memory for these low-level perceptual features (both tone clouds and Gaussian noise). The Emergent Memory Account (discussed in introduction section 1.2A) supports this hierarchical functional recruitment theory of hippocampal function. Further experiments studying neural correlates of encoding and recognition of meaningless stimuli are necessary to fully understand the interaction between subcortical structures, cortical areas and the hippocampus in the perception, storage and recognition of Gaussian noise.

We did not find evidence for involvement of the primary auditory cortex at the group level, neither in the whole-brain analysis nor using an anatomical ROI of Heschl's gyrus. We believe this is due to large inter-individual variations in the anatomical loci

of activated regions. Due to the similarities between the sound types and the presence of all frequency components in each of the sounds, only small activity differences between conditions were detected, and at variable locations for each participant in the primary auditory cortex. The relatively small extent of activated loci with respect to the size of the whole Heschl's gyrus, as well as the inter-subject anatomical differences in location of these loci, might have masked the effect. The involvement of cortical areas in memory for meaningless sounds has been established using the presence of the N1 component in electroencephalography (EEG) measures (Heschl's gyrus, (Andrillon et al., 2015)), as well as using multi-voxel-pattern-analysis (MVPA) computations on fMRI data (planum temporale, (Kumar et al., 2014)). The N1 component (described in chapter 3 introduction) is a negative evoked-potential observed around 100 ms after sound onset in central electrodes. Source-localization of the N1 component using different types of auditory stimuli revealed several cortical sources of this component, all within the Heschl's gyrus (Reite et al., 1994; Zouridakis et al., 1998). This suggests that the physical loci where meaningless acoustic information is stored might vary within the confines of the Heschl's gyrus of each participant. Individual differences in anatomic generators of the N100 component make it difficult to generalize the source across participants. This inter-individual variability (Schönwiesner et al., 2007) would account for the lack of group level activation seen for implicit recognition memory for Gaussian noise in our study.

Interestingly, a lateralization of function was evident from the first synaptic layer in the auditory pathway. The *left* CoN was involved in differentiating old from novel CNs, as well as CNs from Ns. Since the first decussation in the auditory pathway occurs at the

SOC (which performs the function of sound localization in space) the *right* IC receives input from the *bilateral* CoN. Neurons from the IC project to the ipsilateral – *right* – MGB. Our pattern of activation suggests that features are stored hierarchically. Once learned, a simple local inhibitory mechanism - to prevent redundant learning of the same information by multiple neurons – could explain the lateralization observed. That is, even though the stimuli were presented binaurally, once features are encoded by certain neurons, either in the left or right lower auditory pathway, the integrity of hierarchical processing is preserved via local inhibitory mechanisms. Subsequent *storage* of acoustic features possibly occurs in the same neurons that were involved in *perceiving* the stimuli. Support for this claim again comes from STDP models of learning which exhibit such a local inhibitory mechanism allowing neurons to "specialize" detect and encode certain features without redundant learning (Masquelier et al., 2009).While direct comparison of the lateralization of processing during encoding and retrieval is necessary to confirm this hypothesis, activation of the left CoN – right IC – right MGB pathway could suggest that such acoustic information is not equally shared between the lateralized processing pathways. Unlike in vision, multiple decussations exist in the auditory system as highlighted in the introduction (section 1.4A), with some areas along the auditory pathway getting binaural input and other areas getting monaural input. Further experiments explicitly testing the lateralization of recognition memory for Gaussian noise are therefore necessary and one such experiment is described in chapter 5 (overall discussion).

The observed hierarchy of functional specialization along the ascending pathway is also in perfect agreement with predictions from the STDP model (Masquelier et al., 2008,

2009, 2016). This model predicts that output spiking activity from one layer of afferents, when fed into a successive layer, would lead higher-order afferents to have progressively larger receptive fields, and therefore would respond to more complex/ specialized features as seen in our ROI analysis.

3.4.3 Significance of exploratory results

Whole brain exploratory analysis overall yielded very small differences in BOLD activation for all the contrasts of interests. Imaging data from the localizer run contrasting plain noise segments with silence trials showed reliable activations along the auditory pathway. However, the sounds used in the different experimental conditions were perceptually extremely similar and subtle differences in neural processing of these sounds may be hardly detectable using a whole-brain random-effects analysis. This may explain the sparseness of activations observed. Notably, the activity observed was in line with our results from the ROI analysis, with the medial geniculate body being more activated in response to old CNs compared to novel CNs. No regions were evidenced as having a role in detection of cyclicity, again in line with ROI results only implicating the cochlear nucleus in this low-level role.

Cerebellar and posterior cingulate activity was observed in response to old CNs over novel CNs and Ns. Interestingly the cerebellum has been implicated in motor sequence learning and more generally in implicit learning and language (auditory) processing. Overall, the cerebellum seems to be important for accurate temporal computations (Desmond and Fiez, 1998b). Researchers have hypothesized that one role of cerebellar sub-regions is to compare acoustic features representations with the output of sub-vocal articulation online (Desmond et al., 1997). Extrapolating this theory to our results, it is possible that the acoustic features present in incoming sounds are compared with stored features online by cerebellar sub-regions. The increased cerebellar activity for old compared to novel sounds is in agreement with previous findings from a motor

sequence learning study, that demonstrated increasing cerebellar activation with learning under completely implicit learning conditions and without any feedback (Doyon et al., 1996). In other motor sequence learning studies, participants were either given explicit instructions for learning (Friston et al., 1992) or feedback (Jenkins et al., 1994; Jueptner et al., 1997a, 1997b) regarding their response, and a decrease in cerebellar activity was observed with learning. Therefore, the role of the cerebellum in implicit memory might be online internal feedback/error monitoring. In either case, the cerebellar neurons would perform fast, online computations. Theoretical models of non-linear spiking activity in cerebellar Purkinje cells have provided a framework for understanding such computational capabilities (Hakimian et al., 1999). Using a multiplicative probability density function, these authors demonstrated that efficient non-linear dendritic computations result in large fluctuations in output activity along with highly accurate preservation of signal, as seen in Purkinje cell spike outputs in cerebellar nuclei. Therefore, such online comparisons might be the role of the cerebellum in implicit memory for Gaussian noise segments.

Another area that was more activated for old CNs compared to novel CNs was the posterior cingulate cortex (PCC). This region has been consistently involved in memory retrieval processes, with a meta-analysis of fMRI and PET studies demonstrating that the PCC was mostly implicated in long-term memory (Nielsen et al., 2005). Furthermore both the ventral and dorsal subdivisions of the PCC receive extensive projections from the thalamus (Shibata and Yukie, 2003). In fact, the PCC has been implicated in evaluating visual information for emotional content (Vogt et al., 2006) and it is possible that this region performs a similar function for auditory information.

Support for this idea comes from the mapping of projections from the auditory association cortex - linked to delegation of auditory attentional resources – to the posterior cingulate cortex (Yukie, 1995). In fact, internal monitoring, or "attention to internal representations" has been hypothesized to be one potential role of the PCC in the retrieval of memories (Wagner et al., 2005). Accordingly, higher PCC activity in response to old vs. novel CNs could reflect increased attention to existing sound representations. However, the role of the PCC in retrieval has been so far evidenced in *explicit* memory tasks, further experiments are needed to test its precise role in *implicit* retrieval.

3.5 Conclusions and speculations

To conclude, using the frozen noise paradigm in an implicit learning protocol and testing subsequent recognition memory in a fMRI setting, we showed that subcortical regions involved in processing acoustic features also store this information in long-term memory. Notably, the recognition memory for acoustic features seems to depend on participants' sensitivity to acoustic features (a') (experiment 1, figure 2.8) and not on behavioral measures of learning (experiment 2, figure 3.2A). Additionally, this recognition memory is completely implicit, which is consistent with subcortical storage of learned features.

There is also a hierarchy of processing apparent in the subcortical ascending auditory pathway: lower areas like the cochlear nucleus seem to be involved in extracting and processing acoustic features with high temporal resolution, and higher areas like the inferior colliculus and medial geniculate body seem to be involved in detecting more complex ensembles of features. Therefore, while only neurons in the cochlear nucleus were capable of detecting cyclicity, neurons in higher areas were involved in memory for acoustic features. The hippocampus and auditory association cortex are reciprocally connected, a finding that has led to the hypothesis that this connection is crucial for the formation of long-term auditory memories (Kraus and Canlon, 2012). The medial geniculate body and the inferior colliculus project to auditory cortical areas (discussed in detail in the introduction section 1.4A) and it is possible that this IC – MGB - auditory association cortex - hippocampus interaction leads to the consolidation of long-term auditory memories.

Additionally, the cerebellum was also involved in differentiating acoustic features in old and novel sounds, possibly playing a role in auditory working memory comparing features present in incoming sounds to previously stored features and online self-error monitoring of performance. At the cortical level, the hippocampus may have been involved in detecting acoustic features in old sounds suggesting that task requirement rather than stimulus complexity dictates hippocampal recruitment in a task.

While further experiments are required to understand the role of lateralization in storing acoustic features, as suggested by previous behavioral results and the STDP model, we have demonstrated automatic, implicit memory stores along specific nuclei in the ascending subcortical auditory processing pathway.

To summarize, the role of the medial geniculate body in storing acoustic feature information in long-term memory was demonstrated by the results obtained in this experiment. While further experiments are needed to understand the exact mechanism of hippocampus mediated sub-cortical memory for fine acoustic features, these results provide further support for an STDP based encoding mechanism.

IV. Exploring the mechanisms and resolution limits of memory for Gaussian sounds

4.1 Introduction

The finding from the first experiment (chapter 2) demonstrating recognition memory for scrambled versions of learned sounds, even when the scrambling bin sizes were as short as 10-ms, was surprising (figure 2.7). These results were interpreted as participants storing fine acoustic features that are 10 ms or shorter, in sensory memory. Further, short acoustic features unaffected by scrambling at 10 ms bin sizes are implicitly recognized several weeks after learning, suggesting that very short features are stored in long-term memory (Viswanathan et al., 2016). However, an alternate explanation of the finding could be that memory is actually very coarse; that is, very long features, greater than 500 ms in length, would be stored. While scrambling, 10 and 20 ms segments were shuffled randomly (figure 2.2 B). It is important to remember, however, that the post-scrambling distance between any two adjacent segments was always less than 490ms. If neurons encode coarse features longer than 500 ms, it follows that small changes in the temporal order of 10 and 20 ms acoustic segments within the learned feature would not affect recognition memory. Indeed all the segments necessary for effective implicit recognition of a coarse feature still fall within the temporal range enabling recognition from the neurons. Scrambling and looping in that case would not affect implicit recognition memory, as seen in fig 2.7. In other words, if the brain was somehow storing features across the entire 500 ms, it would not matter if the features in the sound were rearranged within this interval.

Results from the previous experiments indicate that subcortical nuclei along the auditory processing pathway, especially the medial geniculate body, could be involved in long-term memory for Gaussian noise (figure 3.4 and 3.6, table 3.2). Response properties of neurons suggest that these neurons carry fine rather than coarse representations (discussed in the introduction section 1.4 A). Additionally, STDP models of learning have been shown to learn short features rapidly (discussed in the introduction section 1.3B, figure 1.10 as well as chapter 2, section 2.5). However, this does not rule out the possibility that other neurons (subcortical or cortical), that are responsible for the observed implicit recognition memory, carry coarse representations of the sound. A major goal of this experiment, therefore, was to test if neurons are really capable of storing fine acoustic features of 10 ms or less. A second question to investigate arises from the behavioral results of experiment 2. Equivalent discrimination of best and worst target CNs, [4 weeks post-learning, figure 3.2 A, and discussed in section 3.4.1] suggests that acoustic features are encoded only based on number of presentations. In that experiment, all CNs were presented the same number of times during both the learning and testing sessions. Number of presentations during the learning session has already been shown to influence learning (Agus et al., 2010), so here, we wanted to ask if number of presentations during the testing session also influences implicit recognition memory.

Based on the results of the first two experiments, another interesting question arises regarding the extent to which stored acoustic information is accessible to conscious processing. In the auditory modality, as in vision, the ventral (what) and dorsal (where) pathways have been identified as processing information consciously and

unconsciously, respectively (Arnott et al., 2004). Therefore, investigating differences in how the same kind of acoustic information is processed when an experimental task requires the engagement of ventral (discrimination task, (Agus et al., 2010)) vs dorsal pathway is an interesting question to pursue, shedding light on how information is processed and handled in the encephalon depending on task requirements. .

As explained above, we aimed at testing several behavioral hypotheses within a single experiment and using an original paradigm. We investigated these hypotheses using three different tasks spanning two experimental sessions.

In the first task, discussed in section 4.3, we tested specific hypotheses regarding the encoding of very short acoustic features using binaural stimuli. In the second task, discussed in section 4.4, we tested our hypothesis regarding the importance of number of presentations during the testing session in long-term implicit recognition memory. In the last task, participants' explicit recognition memory for short acoustic features was tested, and this is discussed in section 4.5. In each task, the specific hypotheses tested are explained along with corresponding analysis, results and interpretation.

4.2 General methods and experimental paradigm

4.2.1 Participants

A total of 12 participants between 20 and 31 years of age participated (mean age = 23.3, S.D = 3.2 years) in the experiment. All participants were compensated for their time with gift cards preloaded with a monetary value of 40 euros. Participants were informed that the purpose of the experiment was to assess the neural correlates of auditory localization and were naïve to the actual hypotheses. All participants gave written informed consent in accordance to the declaration of Helsinki and the University of Toulouse and CNRS requirements for research with human participants [Protocol: CPP14-007a/2013-A01450-45].

4.2.2 Stimuli

Stimuli were programmed and generated using MATLAB R2013 (http://www.mathworks.com/). Each sound stimulus was generated as 2 sequences of 16-bit pseudo-random numbers drawn from a normal distribution with a zero mean and played at a sampling frequency of 44.1 KHz. Such Gaussian sounds have little variation of frequency over time resulting in flat spectrograms. These pseudo-random numbers were generated using seeds that were reset after every trial. Based on trial type, we constructed 2 vectors resulting in Cyclic (CN), Non-Cyclic (N) sequences of equal length. Both vectors were then simultaneously presented to participants in either

ear, to create a sound where a sequence was cyclic in one ear and non-cyclic in the other. This basic trial type is illustrated in figure 4.1A.

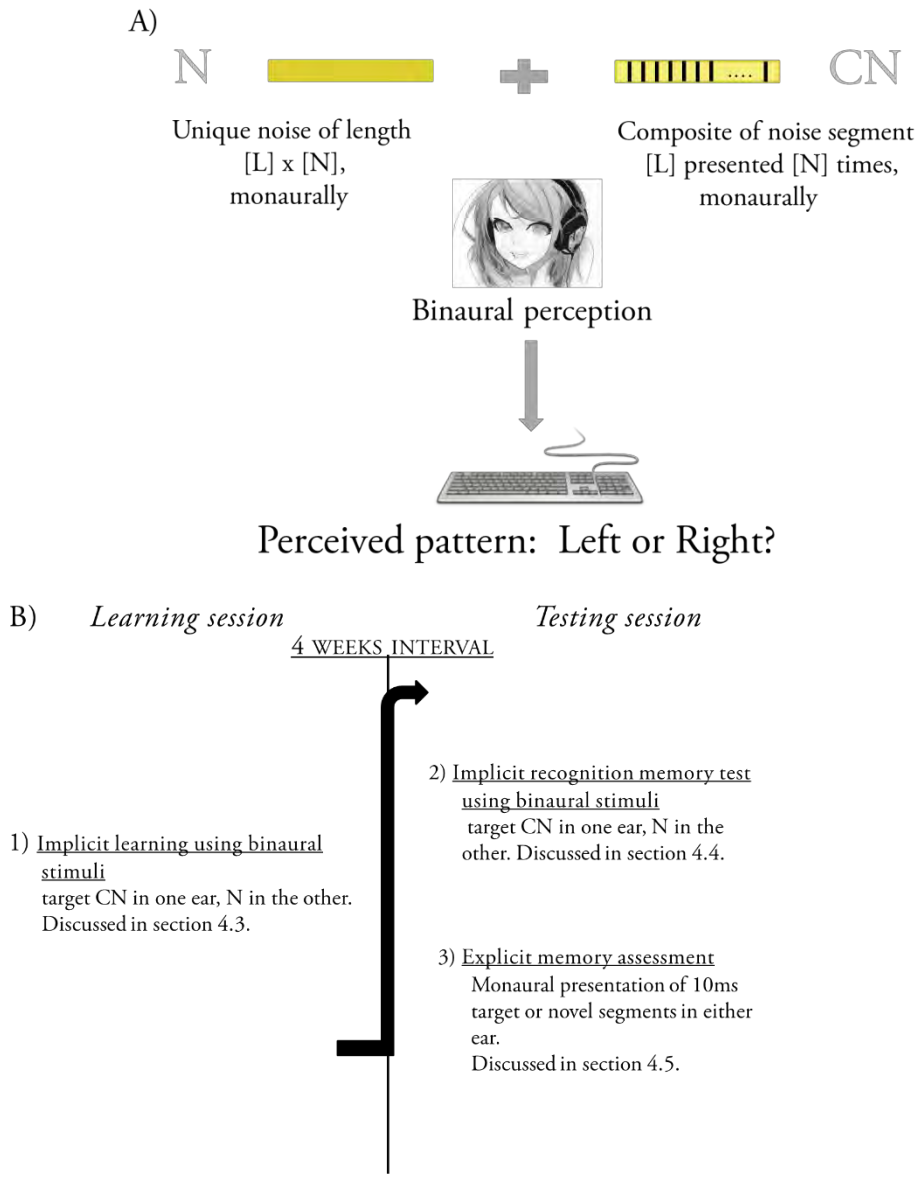


Figure 4.1: This figure provides the general schema of the tasks and trials presented to participants. A) An example of a trial, with a CN in one ear and an equal length N in the other. Participants had to identify the side of the perceived pattern. B) A schematic of both the experimental sessions including all the tasks that participants performed in chronological order.

All the stimuli were normalized to the hearing threshold of the participant, determined from his/her audiogram scores (see procedure below) using the same formula as that used in experiment 2 (chapter 3, equation 2).

As in the auditory discrimination task, unknown to the participant, some exemplar cyclic noises (target cyclic noises) were presented multiple times within a block. Thus, participants were presented with 2 variations of CNs: target and novel CNs. A set of uniquely generated target CNs were presented several times to the participant during the learning and testing sessions. Please note that target CNs will be further referred as old CNs during the testing session (as opposed to novel CNs). A particular target CN was always heard in the same ear. *Novel CNs* and Ns were uniquely generated only heard once throughout the experiment. All stimuli were presented to participants via dual-channel headphones carefully positioned over the EEG cap. Stimuli were delivered using a MATLAB program [MATLAB R2013 (http://www.mathworks.com/)]. The program was also used to record participants' responses. Onset of sound stimulus also sent a trigger to the EEG amplifier as an event marker in the stream of electrophysiological data.

4.2.3 Experimental tasks and procedure

Participants performed four experimental tasks over two experimental sessions, approximately a month apart [mean = 33.25 days, S.D = 4.41 days], as illustrated in figure 4.1B. Each task was designed to test specific aspects of learning, short-term recognition memory (first session) or implicit and explicit long-term recognition memory (second session). Before starting the experiment, all participants were screened

for normal hearing capacity based on the procedure described below. All recruited participants were first screened based on their hearing capacity. The hearing threshold for tones at different frequencies (0.25, 0.5, 1, 2, 4 and 8 KHz) was measured for both ears in an acoustic chamber (designed by studiobricks) using an audiometer (Materiel medical service SARL, France). The hearing threshold was calculated for each ear as in experiment 2 (chapter 3, equation 3). All participants had thresholds at/lower than 20 dB in both ears and were thus included in the study. Once included, all the stimuli for this participant were normalized using equations 2 and 3.

4.2.4 Analysis

Analysis of behavioral data from both sessions was performed using MATLAB and statistical tests were performed using JMP (Version *12.* SAS Institute Inc., Cary, NC, 1989-2007). Analysis of EEG data from both sessions was performed using EEGLAB, an open source toolbox to analyze EEG data (Delorme and Makeig, 2004), in MATLAB.

4.3 Implicit learning of sequences shorter than 10 ms

4.3.1 Hypotheses

Using behavioral measurements, we tested two specific hypotheses regarding encoding of very short noise segments.

3) Implicit memory for acoustic features is fine rather than coarse grained. Therefore, neurons code for features of 10 ms or shorter, and not features greater than 500 ms long.
4) Learning of cyclic noise segments of different lengths will therefore be equivalent. That is, for the same number of exposures, participants will demonstrate similar learning performance for both short and long cyclic noises. This hypothesis is summarized in the top panel of figure 4.3.

4.3.2 Procedure

Noise segments of different lengths, ranging from 10 to 500 ms, were used to create CNs. Participants were presented with 25 back to back presentations of segment lengths, i.e. CNs, in one ear and Ns of equal duration in the other ear. That is, for a segment length of 10 ms, the CN presented in one ear was 250 ms long and an N of 250 ms was presented in the other ear simultaneously. Participants had to indicate the side of the repeating pattern via a keyboard button press. Five different segment lengths

were chosen, which were not multiples of each other – 10 ms, 80 ms, 150 ms, 340 ms and 500 ms. The corresponding CNs were 25 consecutive presentations of these segments resulting in sounds that were 250 ms, 2000 ms, 3750 ms, 8500 ms or 12500 ms long respectively. Examples of a few trials within an experiment block are illustrated in figure 4.2.

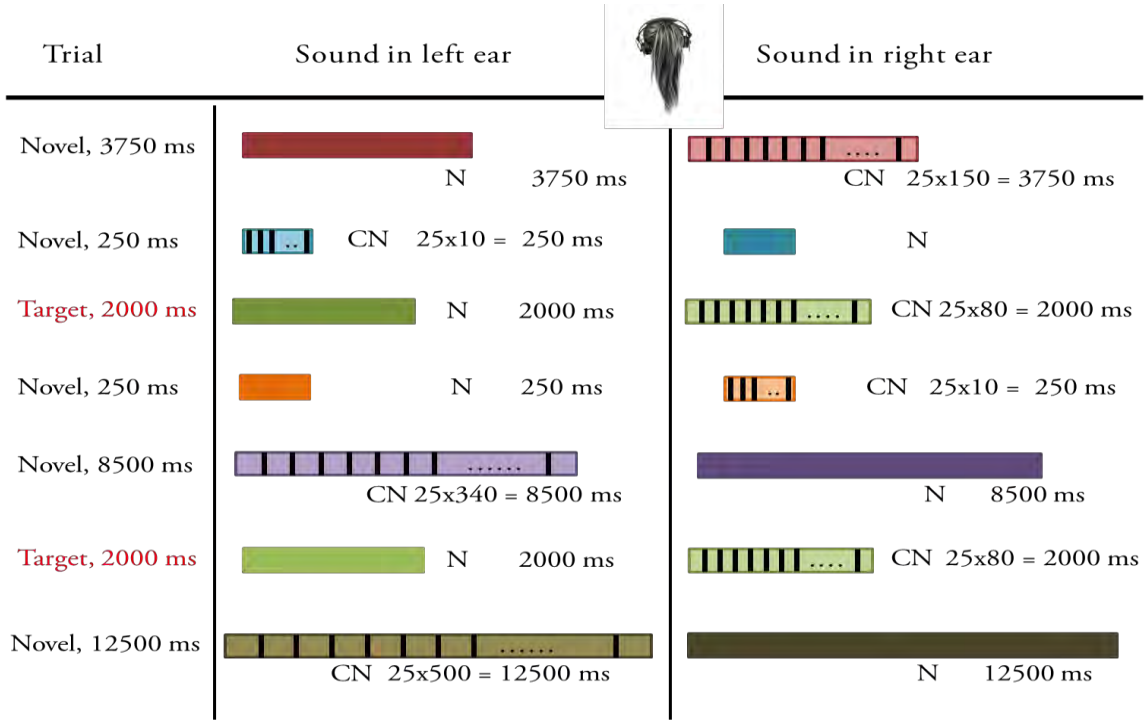


Figure 4.2 Example of a few trials in a block of the implicit learning task. Novel CNs and Ns were all unique but target CN was presented multiple times in the same ear. Trial length varied as a function of size of the segment of noise that was "cycled" 25 times back to back.

Each trial began with participants hearing the CN and then indicating if the perceived cyclicity was on the left or the right. The trials were of variable length and participants did not know how long a trial would last until it started. Once participants had responded, they were presented with the next trial in the block. Participants were given

scheduled breaks between blocks and could take breaks as they preferred within blocks. Participants completed 10 blocks with 50 trials each. Unknown to the participant, we also presented them with *target CNs.* 40 trials in each block were novel – each CN and N was uniquely generated and heard once during the experiment. The other 10 trials were multiple presentations of a *target CN.* A target CN was always presented in the same ear, and the corresponding N was uniquely generated and heard only once throughout the experiment. For each participant, two target CNs were generated for each segment length, with one always presented in the right ear and the other always presented in the left ear. All trials within a block were counterbalanced, with 25 trials with the CN on the left and 25 trials with the CN on the right. The trials were presented in a randomized order and block order (order in which target CNs were presented) was randomized between participants.

4.3.3 Analysis

Following the learning session, the proportions of CNs that were correctly localized was calculated. Localization accuracy for CNs was analyzed as a function of CN segment length as well as trial type (target CN vs. novel CNs).

4.3.4 Results

Overall, participants were able to localize most of the trials (mean = 83.12 % of trials were correctly localized, SD = 0.15), indicating that participants are very good at the task with practically no training. These results are illustrated in figure 4.3, bottom panel.

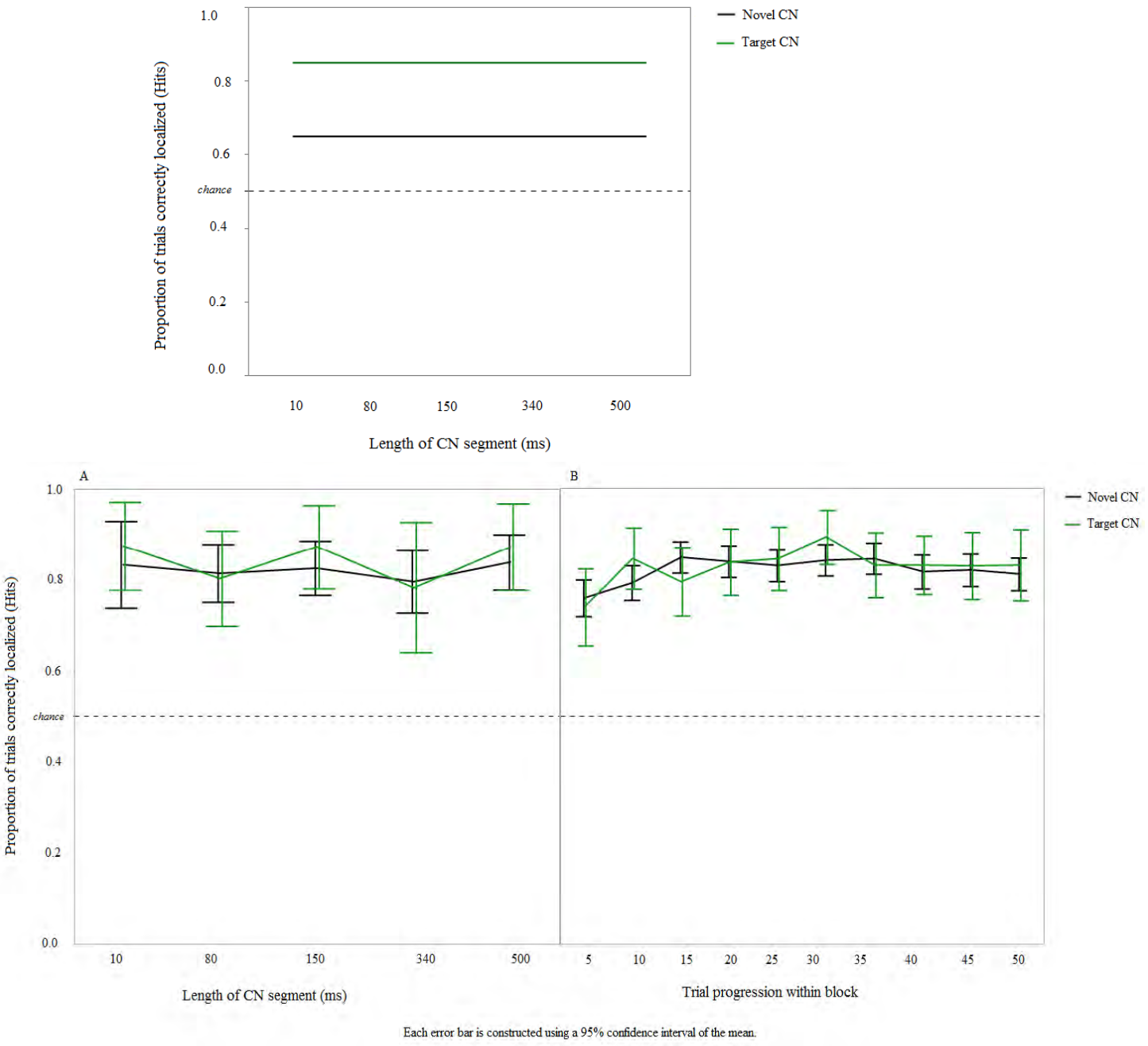


Figure 4.3: Expected and observed localization performance for target CNs vs. novel CNs in the learning session: Top: Hypothesis regarding correct localization of learned and target CNs for different lengths of repeating segment. Bottom: A) Correct localization of learned target and novel CNs for different lengths of the repeating segment (n=12). Participants' localization performance has plateaued by 25 presentations. B) Discrimination rates of target and novel CNs over time (50 trials, mean computed for every 10 non-overlapping trials).

A two-way repeated-measures ANOVA on accurate localization rates was computed using single trials, testing main effects and interaction of within-subjects factors 'trial type' (2 levels, 'target CN' and 'novel CN') and 'CN length' (5 levels, 10, 80, 150, 340 and 500 ms). Neither CN length [$F(4,138) = 2.02$, $p = 0.096$], nor trial type [$F(1,138) = 1.1$, $p=0.296$] were significant predictors of localization performance. The effect of CN length on performance was also equivalent across the two trial types [$F(4,138) = 0.535$, $p=0.71$]. To further investigate the surprising finding that participants localize target and novel CNs equivalently, post-hoc Tukey's Honestly Significant Difference (Tukey's HSD) tests were run. Customized F tests were run for each participant's localization of CNs (test slices), which confirmed that each participant localized target and novel CNs equivalently [all individual p values >0.41]. That is, all participants seem to have reached a performance plateau of localization performance by 25 presentations and did not show any further improvement.

Next, we were interested in the learning of target CNs over time. Another two-way repeated-measures ANOVA on accurate localization rates was computed, testing main effects and interaction of two within-subjects factors. The first factor was 'trial type' (2 levels, 'target CN' and 'novel CN'). The second factor was 'trial progression' (5 levels, average localization over 'trials 1-10', 'trials 11-20', 'trials 21-30', 'trials 31-40' and 'trials 41-50'). Interestingly, we found a significant effect of trial progression [$F(4,5350) = 3.2$, $p = 0.0123$], but not of trial type [$F(1,5350) = 0.37$, $p = 0.542$] and no interaction effects of the two factors [$F(4,5350) = 0.46$, $p = 0.76$]. Using Tukey's HSD tests, the effect of trial progression was further explored. Localization performance remained equivalent throughout the block. These results are summarized in figure 4.3 A and B.

4.3.4 Significance of results

These behavioral results answer several outstanding questions regarding the mechanisms of processing and storing short acoustic features in meaningless noise. Using a novel implicit learning paradigm, we showed that as predicted (hypotheses 1 and 2), memory for acoustic features is fine and not coarse grained. The results are clear when we compare the predicted (figure 4.3 top panel) and actual results (figure 4.3 bottom panel, A). First, localization performance is equivalent and has plateaued (83.12%) for target and novel cyclic sounds, suggesting that 25 presentations are more than sufficient for features in a sound to be learned. It is therefore likely that the (unexpected) equivalent localization performance for target and novel CNs is due to the fact that performance cannot be improved beyond this point. This was also confirmed at the individual level since all participants localized target and novel CNs equivalently, suggesting that by 25 presentations, participants arrive at their 'personal best' for task performance. Additionally, the task is quite easy for participants when compared to the discrimination task used in previous experiments (Agus et al., 2010), where participants were at chance for novel CNs.

Further, performance is equivalent across the different segment lengths, suggesting that neurons code for short features in cyclic noises of variable lengths, resulting in equivalent learning when number of exposures is controlled for. It is possible that participants encode more segments within the longer sounds and therefore these sounds are more *robustly* encoded. However, that cannot be answered using these results. Rather, here we demonstrate that neurons *can* encode features 10-ms long. The

minimum number of presentations after which sounds are learned is a very interesting question which remains to be explored.

Another point to note is that a simple neural mechanism of periodicity detection can explain how participants are able to accurately localize very short, 10ms CNs (since these CNs are presented at a high frequency of 100 Hz). Such frequency detectors have been shown to exist in the auditory cortex to detect repeating features in speech (Kaukoranta et al., 1987). The fact that participants are also able to localize longer length CNs (up to 500-ms used in this experiment) suggests that we might be equipped with low frequency periodicity detectors as well.

4.4 Implicit recognition of very short segments

The aim of this behavioral task was to test the relationship between number of repetitions and implicit recognition memory. Participants performed this task only during the testing (second) session, where long-term memory for sounds learned during the implicit learning task (4.3) was quantified. approximately one month post-learning [mean = 33.25 days, S.D = 4.41 days]).

4.4.1 Hypotheses

Using this task, we tested our hypothesis regarding the importance of number of exposures during the testing session in implicit recognition of previously strongly encoded cyclic noises.

- Performance in long-term implicit recognition memory will increase with number of exposures during the testing session.

.4.4.2 Procedure

A cyclic noise was constructed by concatenating 4, 8 or 16 repeats of a 10-ms noise segment (figure 4.4). This resulted in CNs with durations of 40 ms, 80 ms or 160 ms respectively. The corresponding Ns were constructed as unique noise segments of the CN length. Implicit recognition memory was tested for the two target 10-ms segments that were used during the learning session. In addition, participants were presented with novel CNs that they only heard once throughout this session. All trials were presented in a random order.

Participants were asked to localize the side on which the CN was perceived. Participants performed 72 trials of this localization task with 24 trials for each of the short (40 ms), medium (80 ms) and long (160 ms) conditions. Unknown to the participants, half of the trials in each condition (12 trials, 6 left and 6 right CNs) contained target CNs and the other half contained novel CNs. Note that in during the learning session, segments from recently learned target CNs were presented to participants to quantify short-term recognition memory. During the testing session, segments from target CNs that had been presented 4 weeks previously were used to quantify long-term recognition memory. As in the implicit learning task, all the Ns were uniquely generated. Participants could take breaks between trials as they wished.

The schematic of this task with a few example trials are shown in figure 4.4.

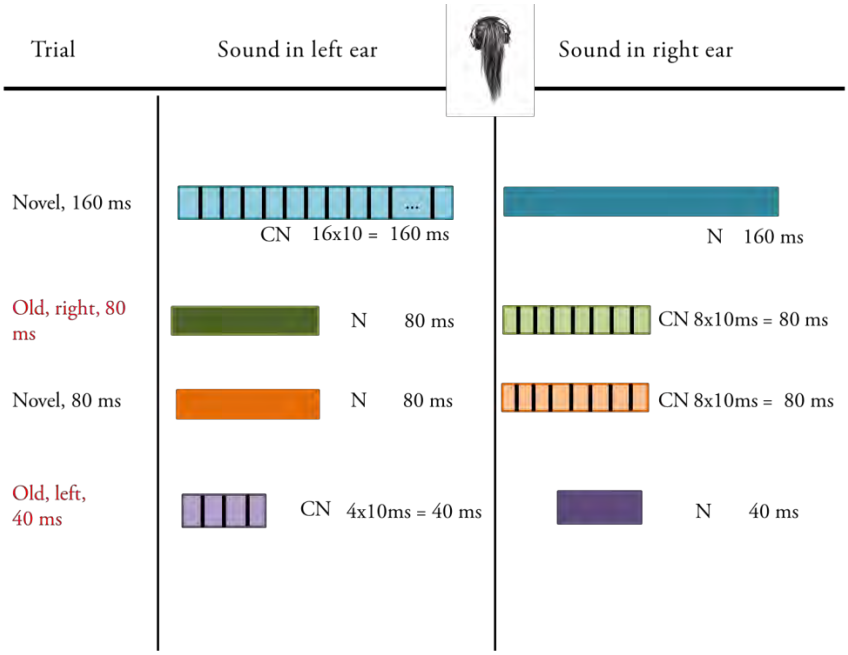


Figure 4.4 Example of a few trials in a block of the implicit recognition task. Novel CNs and Ns were all unique but target CN was presented multiple times in the same ear. Trial length varied as a function of number of presentations of CN.

4.4.3 Analysis

For all 72 trials, the proportion of hits was calculated. The rate of accurate localization of CNs was analyzed as a function of CN length and trial type (old CNs vs. novel CNs).

4.4.4 Results

The results of this task are summarized in figure 4.5, bottom panel.

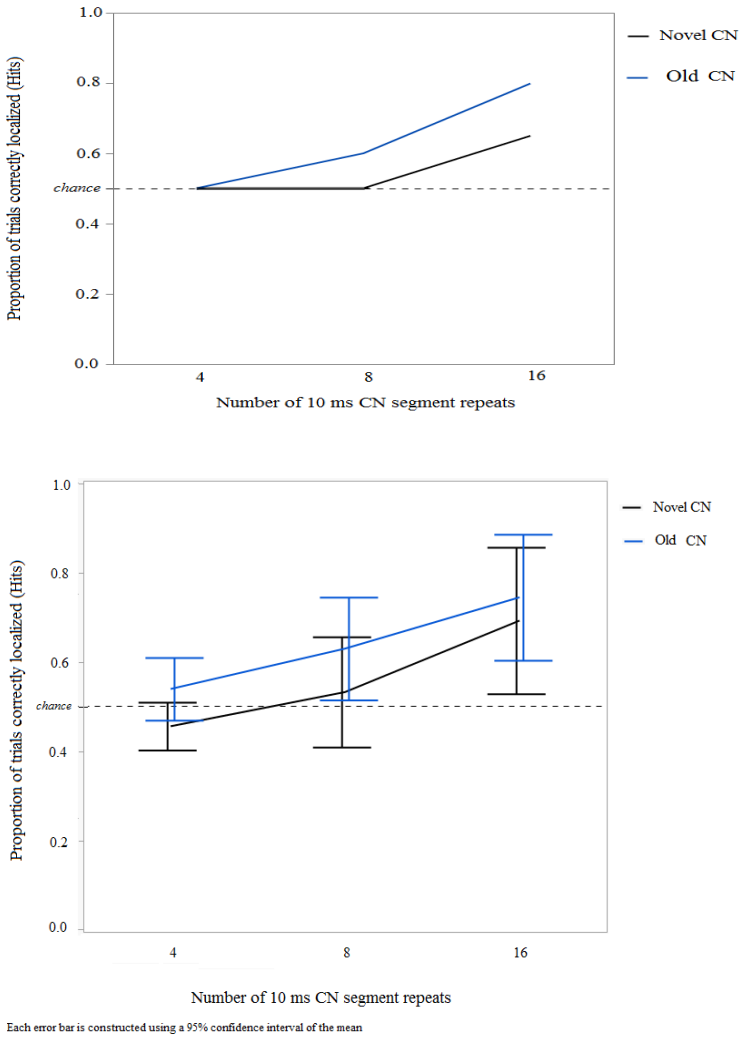


Figure 4.5 Expected and observed localization performance for old (target) CNs vs. novel CNs in the implicit recognition task, for increasing number of repeats of the 10-ms segment. Top: Predicted localization performance for different number of repeats. Bottom: Observed localization performance as a function of number of repeats (n=12).

A two-way repeated-measures ANOVA on localization performance tested main effects and interaction of within-subjects factors 'trial type' (2 levels, 'target CN' and 'novel CN') and 'segment repeats' (3 levels, 4, 8 and 16). Interestingly, both the trial type [$F(1,72) = 5.29$, $p = 0.0252$], and number of segment repeats [$F(2,72) = 14.55$, $p<0.0001$] were significant predictors of localization performance. However, there was no interaction between the two factors [$F(2,72) = 0.155$, $p=0.855$]. These results imply that localization of learned target CNs was more precise than localization of novel CNs, irrespective of the number of times the 10-ms segment was presented, overall. To further investigate this, one-sample t-tests were run for the localization performance for each of the trial types for each number of segment repeats against a theoretical mean of 0.5 (null hypothesis that participants are at chance at localizing the sounds). Interestingly, at 4 segment repeats (very short sounds), localization was at chance for both target [$p=0.12$] and novel CNs [$p=0.19$]. However, at 8 segment repeats, while localization was at chance for novel CNs [$p = 0.29$], target CNs were localized significantly better than at chance [mean = 0.579, SE = 0.028, $t(288) = 3.879$, $p <0.006$ (value corrected for multiple comparisons)]. At 16 segments repeats, localization was significantly better than chance for both novel [mean = 0.691, SE = 0.027, $t(288) = 7.001$, $p <0.006$ (value corrected for multiple comparisons)] and target [mean = 0.743, SE = 0.026, $t(288) = 9.424$, $p <0.006$ (value corrected for multiple comparisons)] CNs.

To understand the significant effect of number of repeats, Tukey's Honestly Significant Difference (Tukey's HSD) tests were used. This test showed that localization performance was significantly better for 16 repeats than for both 4 [effect size ($mean_{(i)}$-$mean_{(j)}$) = 0.22, $CI_{95\%}$ = (0.12, 0.32), $p < 0.0001$] and 8 [effect size = 0.14, $CI_{95\%}$ = (0.04,

0.24), p = 0.0045] repeats. Notably, the localization performance was equivalent when the 10-ms segment was presented 4 or 8 times [p = 0.117].

4.4.5 Significance of results

Implicit long-term recognition for learned features was demonstrated by a clear memory effect in localizing target CNs over novel CNs. This memory effect was evident even for short sounds of 80 ms, i.e., participants were able to localize 80 ms target CNs better than chance, while they were still at chance for same length novel CNs. As we predicted, localization is at chance for very short sounds (4 repeats), and the memory effect is only apparent with 8 repeats. We also observed that localization improved with number of presentations for both old and novel sounds, indicating that the number of presentations critically affects behavior. As in the learning session, there was a behaviorally measurable effect of number of repeats on localization performance for novel CNs. Taken together, these results imply that between 8 to 16 presentations, novel CNs are reliably localized and between 16 to 25 presentations, localization performance reaches a plateau, demonstrating an 'online' learning curve for novel CNs.

4.5 Explicit recognition of short segments of noise

The aim of this behavioral task was to test participants' explicit recognition memory for noise. Participants performed this task only at the end of the testing (second) session.

4.5.1 Hypotheses

Using this task, we tested the following hypothesis:

- Acoustic information encoded using a task engaging the dorsal auditory processing pathway, such as a sound localization task, would be inaccessible to conscious processing. Participants will therefore demonstrate implicit recognition memory but not explicit recognition memory for these meaningless features.

4.5.2 Procedure

In the last task, explicit memory was tested for old CNs (target CNs presented during the learning session). To avoid confounds of recent reactivation of learned sounds during the previous tasks, target CN segment of 150 ms length from the first session, both left and right, were used. These target CNs were not used in any other task during this session. To perform the task, participants were presented with novel noise segments of 150 ms length as well as target/old target CN segments of 150 ms and asked to explicitly identify if they have heard the sound previously (old vs new discrimination task). Please note that old target segments were always presented in the

same ear as they had been presented during learning. All the sounds were presented monaurally, with silence in the contralateral ear for each trial.

4.5.3 Analysis

Participants' classification of a sound as 'old' or 'novel' was analyzed in terms of actual trial type. This analysis was performed at both the individual and group level to identify individual participants capable of distinguishing old from novel noise segments as well as any group level tendencies.

4.5.4 Results

The results from this experiment are illustrated in figure 4.6.

Analysis of participant responses from the explicit memory task showed that trial type (2 levels, 'learned' and 'novel') had no effect on participants' correct identification of learned sounds [$F(1,480) = 0.36$, $p = 0.549$]. We were also interested in seeing if any of our participants were able to successfully differentiate learned and novel CNs at the individual level.

Lastly, as predicted, no participant demonstrated explicit recognition memory in a task which engaged areas in the dorsal (unconscious) processing pathway.

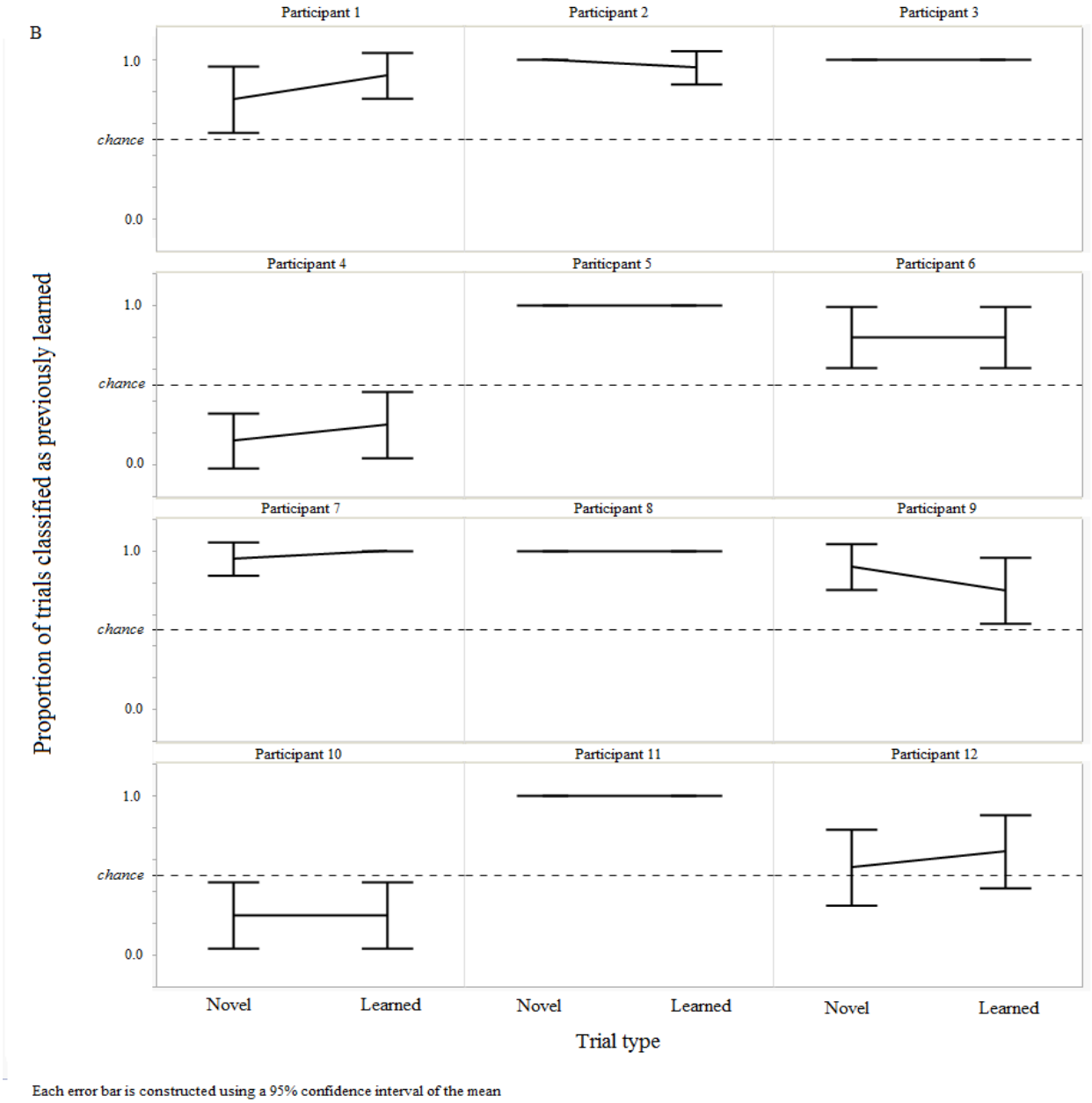


Figure 4.6: Individual performance in the explicit recognition memory task during the testing session. The proportion of trials classified as learned by each participant for both learned and novel CNs is shown here. Error bars are 95% confidence intervals of the mean (n=12).

4.6 Conclusions and speculations

Before launching into a discussion of the implications of these findings, I would like to mention the task and the implications of monaural learning. First, while piloting the experiments, we observed that there is an ambiguity in localization of the cyclic noise with smaller number of repeats, as is evident from figure 4.5(bottom). That is, with fewer repeats, participants could tell that something was repeating but were unsure of the side. This ambiguity is completely resolved by 25 presentations (seen in figure 4.3 bottom A). How such computations are made in the brain are fascinating when we consider that with the exception of the first synapse (the cochlear nucleus) all nuclei receive binaural information (introduction section 1.4A). Thus resolving such ambiguity needs to be further studied to understand how these nuclei "untangle" information coming from both ears.

Analyzing the behavioral results tells us that recognition memory for acoustic features seems to depend on participants' sensitivity to acoustic features (a') (experiment 1, figure 2.8) as well as the number of exposures (experiment 3, figure 4.5bottom) and not on behavioral measures of learning (experiment 2, figure 3.2A). Also, as seen in experiment 2, participants do not seem to have conscious access to stored information, perhaps since fine features might be stored at the subcortical level, areas that do not participate in conscious processing.

These results highlight the importance of number of exposures in dictating performance. During the learning session, it was observed that 25 presentations of a

segment are enough for learning and during the testing session, it was observed that even with 8 presentations, participants performed better than chance for target but not for novel CNs (memory effect). With novel sounds, we showed that (i) with 8 presentations, localization performance is at chance, (ii) with 16 presentations, participants start reliably localizing sounds and (iii) by 25 presentations participants had reached a performance plateau. These findings answer key questions regarding the relationship between number of presentations and learning and retrieval. A potential way to further investigate such questions would be to modify the experimental paradigm described in this experiment during the learning session: by presenting the noise segments (target and novel) a variable number of times during learning, the strength of memory can be correlated to the strength of encoding. At any rate, this experimental paradigm, like that designed by Agus and colleagues (Agus et al., 2010) can be used to answer a plethora of questions regarding the characteristics of implicit memory for Gaussian noise. Combining these results we see that behaviorally, a memory effect is present in long-term memory (evident with 8 repeats, figure 4.5, bottom panel). It is thus apparent that neurons are capable of learning fine temporal features, 10 ms or less, robustly, and store this information in long-term memory. These results, like the results of the previous experiment, are in line with predictions from an STDP based hierarchical learning mechanism.

Lastly, there are clearly both conscious and unconscious aspects of detecting and encoding acoustic features present in noise. While participants are able to hear cyclicity consciously, they are unable to have conscious access to individual feature information. A similar result was observed here.

V. CONCLUSIONS AND PERSPECTIVES

GENERAL CONCLUSIONS

The experiments described here and conducted as a part of my thesis have been aimed at testing specific hypotheses regarding of how sensory information is processed and stored, and thereby improved our understanding of these mechanisms. The principal findings and their implications are discussed below.

1) Individual acoustic features that are as short as 10 ms are robustly stored.

Behavioral results from chapter 2 (figure 2.7) and 3 (figure 3.4A) show that memory for auditory meaningless stimuli involves storing extremely short features in noise. Since auditory functions such as localization of sound sources in space necessarily involve high resolution acoustic perception, the finding that memory for acoustic features is equally high resolution suggests that perception and memory share anatomical resources, as discussed in the introduction section 1.1C. This is in line with predictions of the emergent memory account (discussed in introduction section 1.2A), which says that memory is non-modular and that rather, task complexity and computational requirements determine what areas are recruited for perception *and* memory. This makes sense from an evolutionary perspective: since the brain has evolved to be highly

efficient and energy-cost effective (Achard and Bullmore, 2007), processes that share computational demands would naturally share neural resources[9].

One factor that possibly influences strength of recognition memory for a given sound is the number of features that are stored. If several patterns within a sound are stored, disrupting a fraction of these (via scrambling, for instance), would be less likely to disrupt implicit recognition. This would explain how participants with higher a' values during implicit learning have higher implicit recognition performance for intact, looped and scrambled versions of target CNs compared to novel CNs [for which participants are at chance level, (figure 2.8)] – participants with higher sensitivity to acoustic features, resulting in better discrimination performance, might be storing several non-overlapping features in each target CN. This leads to several follow-up interesting questions like: does the brain learn all repeating patterns? What other factors influence how many features are stored in a target CN? What is the relationship between number of non-overlapping features that are stored and robustness of recognition memory to scrambling and other transformations of the learned stimuli and does this relationship plateau at some point? In other words, is there an 'optimal' number of features to be learned in a stimulus which would lead to robust encoding without wasting neural resources by redundant storage? The latter idea is particularly interesting since it also raises theoretical questions regarding brain function. On the one hand, one can argue that all repeating patterns might be stored purely because of repetition. STDP models of

[9] Although not traditionally discussed in this context, in my opinion, support for this idea comes from research on face processing. Using an fMRI study, it was demonstrated that bird experts and car experts recruited the brain regions involved in face processing (notably the fusiform face area) when making identity judgements regarding birds and cars respectively (Gauthier et al., 2000). Since the brain didn't evolve to differentiate cars, it is likely that these regions were recruited to the task based on similarity of computational requirements.

learning noise patterns as well as our experimental data support this claim to a certain extent (discussion of chapter 3, section 3.4.1 as well as results of chapter 3, figure 4.3 bottom panel). The assumption then is that the brain considers any repeating stimulus *potentially* important and stores them in anticipation of associated behavioral significance. On the other hand, storing *all* repeating input indiscriminately and indefinitely would be an inefficient strategy. Researchers have just started scratching the surface of topics like how decisions are made in the brain regarding acoustic feature processing (what and how many) and subsequent storage (how long) and several interesting hypotheses need to be tested to answer these questions. One factor that might influence such computations is attention and further experimentation is required to understand exactly how attention impacts detection and storage of acoustic features.

These results are in complete agreement with predictions from STDP models. However, some behavioral aspects are not yet accounted for by these models. One example of this is the finding that participants with high a' values during encoding perform better on tests of implicit recognition memory. While STDP model neurons are able to learn acoustic features with high selectivity (which can then be subsequently labelled), what it currently doesn't explain is the idea of individual variability in participants' ability to do the task. Inclusion of additional parameters in computational STDP models, such as factors accounting for pre-existing biases that allow certain features to be encoded by some participants but not others might be necessary to account for such findings. At its core though, STDP is a likely candidate mechanism explaining how statistical regularities are implicitly extracted and stored in layers of neurons.

To summarize, results from these experiments show that implicit recognition memory for Gaussian noise is robust, highly specific, fine grained and resistant to transformation and interference. The specificity of memory is evident for very short acoustic features that are 10 ms or shorter. All these findings are in agreement with predictions from STDP models of learning repeating patterns in noise, implicating STDP as the candidate cellular mechanism for storing sensory information.

2) Participants do not show any explicit memory or conscious access to these stored acoustic features.

Behavioral results from these experiments, principally from the explicit memory tasks used in experiments 2 and 3, show that acoustic features are robustly stored without any conscious access. Not a single participant we tested in experiment 2 (n=15) or experiment 3 (n=12) was able to reliably consciously differentiate 'old' from 'novel' Gaussian sounds. Interestingly, this lack of conscious access was even observed in tasks that tap into conscious processing networks. As discussed in the introduction section 1.4A, cortical processing of acoustic information occurs via the dorsal and ventral processing pathways, similar to vision (figure 1.18). The dorsal pathway has been implicated in non-conscious computations involving orientation towards sound objects in space and is referred to as the 'where' pathway. The ventral pathway, on the other hand, has been implicated in conscious computations involving identifying sound objects in our environment and is referred to as the 'what' pathway (Arnott et al., 2004). By using tasks that need participants to make either identity judgements [cyclic/non-

cyclic discrimination, experiment 1 and 2] or localization judgements [left/right source of pattern discrimination, experiment 3] we were able to test conscious access to information stored when acoustic information is processed in either pathway. Our results indicate that participants were unable to consciously identify target sounds in either task. The same sounds had been successfully encoded in both tasks; participants demonstrated implicit recognition for the same sounds they couldn't identify explicitly. These results suggest that the cortical processing pathway does not influence the lack of explicit recognition of stored acoustic features.

Perhaps this complete lack of conscious access to acoustic features can be attributed to subcortical storage of information. Our neuroimaging data certainly support this idea (figure 3.4), which is further discussed under the implications of subcortical storage of acoustic information. Another reason for the lack of conscious access to such information might be that such features are stored without associations or labels. While repeating features are clearly noticeable in cyclic noise, the mechanism of encoding acoustic features via *selectivity* doesn't result in conscious perception of individual features. In this hypothetical mechanism, neurons encoding acoustic features might 'silence' surrounding neurons via intra-cortical inhibition. Novel and target CNs would then be differentiated purely based on activity differences across a layer of neurons, with previously encoded features in target CNs inducing localized activity of neurons selective to the feature surrounded by a silent zone, and novel CNs inducing diffuse activation across neurons in the layer. This mechanism is in line with sparse coding mechanisms for storing information as discussed in the introduction section 1.2A. These are fascinating ideas that need to be further explored.

3) *Some of the factors affecting implicit recognition memory for meaningless acoustic features have been identified.*

When considering the behavioral results from all three experiments together, some of the factors affecting implicit memory become clear. First, implicit recognition memory seems to depend on measures of participants' sensitivity to differences between individual acoustic features (a') and not on measures of learning for a specific sound (classifying as "best" or "worst" CN) (figure 2.8, figure 3.2A). In fact, results from regression to the mean analyses indicate that participants learn features in sounds irrespective of conscious perception of cyclicity (figure 3.3). Further, a strong link between number of presentations and efficiency of recognition was observed (figure 4.5, bottom panel). Therefore, it seems that implicit recognition memory depends on number of presentations and participant's sensitivity to detect acoustic features but not on other, experimenter defined measures of their behavior.

However, factors that determine a participant's sensitivity (a') are still unclear. It is likely that sleep, prior musical training and aptitude, and auditory imagery capabilities affect a' values, although, as evident from our data, more objective measures of these factors are necessary.

The relationship between attention and implicit recognition memory for Gaussian sounds is also fascinating. Attentional networks do not sample the world continuously, and how fluctuations in attention dictate what acoustic features are processed and

stored remains fascinating but unclear. Especially interesting to study is the interaction between arousal, bottom up attention and encoding mechanisms. One possible way of investigating this is by understanding the activity of the Locus Coeruleus (LC) with respect to implicit recognition memory. The LC, a brainstem nucleus part of the metencephalon, has been recently implicated in memory (Jacobs et al., 2015) and may be the very first location of intra-cellular lesions in Alzheimer's disease (Braak et al., 2011). The primary function of the LC is to modulate arousal by releasing norepinephrine and thereby controlling the engagement of the sympathetic nervous system (Aston-Jones et al., 1991, 1996, 1999). Putting these results together, it follows that attention dictates strength of encoding and therefore retention by mediating bottom up attention and attentional cycles. The LC could also modulate implicit recognition memory via interactions with the hippocampus [seen to be involved in memory for meaningless auditory sounds (Kumar et al., 2014) and figure 3.5], independent of attentional modulation. Norepinephrine has been demonstrated to modulate hippocampal long term potentiation (Hopkins and Johnston, 1984; Stanton and Sarvey, 1985, 1987) and ascending fibers from the LC project to the hippocampus (Jones and Moore, 1977; Pickel et al., 1974). LC activity, which releases norepinephrine in the hippocampus, could modulate hippocampal-cortical interactions. As discussed in introduction section 1.2A, these hippocampal-cortical interactions help store information in long-term memory. Therefore, by precise temporal modulation of hippocampal LTP activity, norepinephrine released from LC activity can help store acoustic features in meaningless features in LTM.

Monitoring the level of LC engagement during encoding and implicit recognition would allow us to test these hypotheses. Since the release of norepinephrine exclusively controls pupil dilation (a function of the sympathetic nervous system), using pupillometry to track the modulation of LC activity (in real time) and thereby arousal is one strategy for testing such hypotheses.

4) Acoustic features are stored at the subcortical level, with the medial geniculate body playing a role in long term storage.

Behavioral data from these experiments, in agreement with predictions from STDP models, suggest that feature extraction and storage can occur in early auditory areas. Results from our fMRI experiment also support this idea, a finding that was interpreted as these different subcortical regions performing various roles in feature processing and storage.

First, it appears that the medial geniculate body is implicated in long-term memory for acoustic feature information (figure 3.4 and 3.6, table 3.2). Since there are efferent connections from MGB to the lower areas such as the inferior colliculus and the cochlear nucleus, it is possible that via top down modulations, specific neurons in these regions might also be able to differentiate recent memories and remote memories, an intriguing idea that needs to be tested using future experiments. The implication of both MGB and primary auditory cortex in storing acoustic information (Andrillon et al., 2015; Kumar et al., 2014), along with the agreement of our data with predictions from STDP

models suggest that sounds are processed hierarchically along the auditory pathway. This hierarchy of processing progressively specialized features is also evident at the cortical level since Luo and colleagues demonstrated that cortical sources show 'slow' phase tracking of learning individual sounds: this phase tracking to individual target CNs was observed only 500 ms after sound onset (Luo et al., 2013). These results demonstrate the importance of precise temporal measurements that allow characterization of the feed-forward information transfer along the ascending auditory pathway. In other words, without precise temporal markers such as the N1 (seen around 100 ms after sound onset) which are clearly part of the feed-forward pathway, we cannot address hypotheses regarding the actual storage of information along the lower auditory pathways. On the other hand, how top down modulations affect perception and memory [discussed in introduction section 1.4A and 1.4B], are also fascinating to investigate and using precise spatial measurements allows the identification of all the areas involved in processing and storing. To illustrate this using an example, while the influence of efferent connections from the MGB might influence processing of auditory stimuli in lower areas like the inferior colliculus, temporal measurements alone will not be able to answer this question. Therefore, using a combination of techniques to investigate hypotheses (as we have tried to do over the experiments described in this thesis) regarding memory would result in a more comprehensive understanding of underlying mechanisms than any stand-alone method would.

If, as suggested by our results, acoustic features are indeed stored at the subcortical level, it could be one reason for the lack of conscious access to individual feature

information, since the thalamus can act as a gate for information transfer from and to the cortex, a function that is especially evident from research on thalamic functions during sleep (Steriade et al., 1990, 1993; Tsoukatos et al., 1997). Thalamic oscillations regulate sensory input into the cortex during sleep, allowing interference–free consolidation of information stored in cortical areas. Therefore, an interesting consequence of storing information at the subcortical level would be the higher interference from incoming sensory input, including during sleep, since these areas receive sensory input constantly. How feature information is stored in the long term remains unclear except for the role of descending modulation from cortical areas. This idea is also supported by the conclusions of Andrillon and colleagues that idiosyncratic features of Gaussian noise are stored in the primary auditory cortex, as evidenced by the modulation of the N1 component amplitude with learning (Andrillon et al., 2015). Perhaps consolidation and storage of temporally precise segments is mediated by the hippocampus [(Kumar et al., 2014), and the results discussed in chapter 3, figure 3.5], but is physically stored in subcortical areas with high temporal resolution for individual idiosyncratic features present in target CNs. Thalamo-cortical interactions, cortico-fugal influences, bottom up attention as well as arousal probably modulate the quantity, quality and efficiency of this storage. The precise mechanism of this hypothetical *subcortical consolidation* and the role of sleep in selective pruning of subcortically stored information need further experimentation to be understood.

FUTURE PERSPECTIVES

As demonstrated by the experiments preformed in this thesis, using Gaussian noise to investigate memory mechanism is very interesting. The capacity for generating novel stimuli is pretty much inexhaustible, with extremely low probabilities of randomly generating similar sequences. Therefore, using elegantly designed experimental paradigms [like the implicit learning cyclic/non cyclic discrimination task used by Agus and colleagues (Agus et al., 2010), or the implicit learning sound localization task explained in chapter 4] will allow us to continue answering a plethora of questions regarding the mechanisms of processing and storing meaningless auditory features. Besides the ideas discussed above, several future experiments are possible

The novel experimental paradigm described in chapter 4, for instance, can be used to understand the relationship between number of presentations and learning of acoustic patterns, by varying the number of presentations during the learning session and then testing subsequent implicit recognition 4 weeks later. This paradigm can also be used to further understand the lateralization of function observed along the auditory processing pathway. Using monaural presentation of target CNs, we can investigate what happens when a sound presented in one ear and subsequently encoded along the lateralized processing pathway during learning, is later used to test *implicit recognition in the other ear* and is processed by the contralateral ascending pathway. Using this paradigm in an fMRI experimental setting could answer questions regarding exactly

how information is stored along the ascending pathway. In fact, we have run pilot participants to test this question and to optimize the localization task for the scanner.

Another possible series of experiments with the potential to answer several questions regarding the spatial and temporal correlates of memory is to run implicit learning and memory tests using a more specialized system to collect the auditory brainstem response. Further, despite our efforts to address the question, the roles of musical expertise, sleep and attention on encoding and implicit recognition remain unclear. Experiments designed to address such questions need to be developed. As an example, perhaps using a divided attention paradigm or directly measuring sleep parameters that might influence memory, while conducting implicit memory tests, will help us understand the influence of these parameters on implicit recognition memory.

Last, but not least, while these stimuli are great to study memory in a lab setting and they help understand memory, they can also be used to understand how features in natural sounds are encoded and stored. Several sounds in nature are intrinsically periodic and understanding memory for cyclic noises will help understand how natural cyclic signals in sounds are processed and stored. Examples of such naturally occurring sounds are bird calls and animal sounds such as toads. At the outset, bird calls from the same species might all sound alike. However, even lay listeners without training are able to detect unique features in individual sound samples. With training and enthusiasm, people can differentiate and identify highly similar calls, as demonstrated by the capabilities of bird experts. This ability to detect fine acoustic features is not so

different from abilities that trained/skilled musicians have to differentiate subtle changes in melody and composition of highly similar pieces of music.

Using animal and bird calls to investigate perception and memory for naturalistic auditory stimuli has two advantages. Firstly, it allows us to investigate at what point semantic labels are added to such stimuli, a feat which is almost impossible with Gaussian sounds. While features present in these stimuli fall in a narrower frequency and spectral range than Gaussian sounds, this makes individual exemplars perhaps easier to identify and therefore questions regarding identifying and labelling a sound can be addressed. Secondly, a preference in processing auditory features present in a certain range of harmonics has been demonstrated in humans, a finding that has been interpreted as being caused by fluctuations within this range being the most biologically relevant (Gill and Purves, 2009). This range includes all the features present in human vocalizations as well as several preferred music scales. Therefore, memory for acoustic features within this preferred range may be even more specific, robust and high resolution, a hypothesis that is better tested using stimuli with a narrower range of features, such as natural animal and bird vocalizations.

Going back full circle, perhaps the ability to make fine discriminations regarding such naturally occurring cyclic stimuli was one of the reasons, along with sound localization, that the auditory system evolved to store such temporally precise information. A whole avenue of research remains to be explored comparing how information is sampled and processed in different domains. An attempt to have participants learn visual noise (Gold et al., 2014) has already revealed that while participants can encode meaningless

visual sequences similar to audition, the memory for mirror-reversed sequences is quite poor, suggesting that the resolution of features that are stored is different in the visual domain. Humans are able to discriminate meaningless textures and tactile patterns (Connor and Johnson, 1992; Lederman and Klatzky, 2009), but long term memory and mechanisms of this processing remain to be explored as well. Interestingly, processing along the ascending somatosensory pathway is strikingly similar to processing along the ascending auditory pathway. For instance, both lateralization of function and high temporal accuracy in judgements of pattern similarities have been reported in the tactile domain (Blake et al., 1997; Craig and Baihua, 1990). Therefore, comparison of perception and long-term memory for feature information across these domains could lead to insight into how these systems evolved as well as the kind of computations that are important for both. Relatively little is known about memory for meaningless stimuli in other sensory modalities. The ability to make fine discriminations in taste and smell would help survival and are thus possible sources of selective pressure in evolution. In this case, the resolution for sensory storage might be specific and fine grained in these modalities, but can 'meaningless' odors and tastes be learned? Given the strong link between perception and emotion in these modalities, it is an interesting idea to pursue in future experiments.

In the experiments conducted during my thesis, by testing specific hypotheses regarding auditory sensory memory using a combination of behavioral, imaging and computational techniques, and by using both well established and novel experimental paradigms, we have demonstrated the robustness and mechanisms of memory. The main take away seems to be that fine acoustic features, as short as 10 ms, are stored

robustly, and subcortical regions such as the medial geniculate body are involved in storing these features in long-term memory. Just as we have been inspired by previous experiments to ask novel questions, I hope these experiments and findings inspire future experiments into understanding how such information is stored and the factors which influence this process.

VI. REFERENCES

Abraham, W. C. (2003). How long will long-term potentiation last? *Philos. Trans. R. Soc. London B Biol. Sci.* 358.

Achard, S., and Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comput. Biol.* 3, 0174–0183. doi:10.1371/journal.pcbi.0030017.

Aertsen, A. M. H. J., and Johannesma, P. I. M. (1981). The Spectro-Temporal Receptive Field. *Biol. Cybern.* 42, 133–143. doi:10.1007/BF00336731.

Agus, T. R., and Pressnitzer, D. (2013). The detection of repetitions in noise before and after perceptual learning. *J. Acoust. Soc. Am.* 134. doi:10.1121/1.4807641.

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

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, 53–64. doi:10.1007/BF00229109.

Alain, C., Woods, D. L., and Knight, R. T. (1998). A distributed cortical network for auditory sensory memory in humans. 23–37.

Alberini, C. M. (2007). Reconsolidation: the samsara of memory consolidation. *Debates Neurosci.* 1, 17–24. doi:10.1007/s11559-007-9000-z.

Amaro, E., Williams, S. C. R., Shergill, S. S., Fu, C. H. Y., Macsweeney, M., Picchioni, M. M., et al. (2002). Acoustic Noise and Functional Magnetic Resonance Imaging : Current Strategies and Future Prospects. 510, 497–510. doi:10.1002/jmri.10186.

Aminoff, E. M., and Tarr, M. J. (2015). Associative Processing Is Inherent in Scene Perception. *PLoS One* 10, e0128840. doi:10.1371/journal.pone.0128840.

Anderson, L. A., and Linden, J. F. (2011). Physiological differences between histologically defined subdivisions in the mouse auditory thalamus. *Hear. Res.* 274, 48–60. doi:10.1016/j.heares.2010.12.016.

Andrillon, T., Kouider, S., Agus, T., and Pressnitzer, D. (2015). Perceptual Learning of Acoustic Noise Generates Memory-Evoked Potentials. *Curr. Biol.* 25, 2823–2829. doi:10.1016/j.cub.2015.09.027.

Arnott, S. R., Binns, M. A., Grady, C. L., and Alain, C. (2004). Assessing the auditory dual-pathway model in humans. 22, 401–408.

doi:10.1016/j.neuroimage.2004.01.014.

Aston-Jones, G., Chiang, C., and Alexinsky, T. (1991). "Chapter 35 – Discharge of noradrenergic locus coeruleus neurons in behaving rats and monkeys suggests a role in vigilance," in *Progress in Brain Research*, 501–520. doi:10.1016/S0079-6123(08)63830-3.

Aston-Jones, G., Rajkowski, J., and Cohen, J. (1999). Role of locus coeruleus in attention and behavioral flexibility. *Biol. Psychiatry* 46, 1309–1320. doi:10.1016/S0006-3223(99)00140-7.

Aston-Jones, G., Rajkowski, J., Kubiak, P., Valentino, R. J., and Shipley, M. T. (1996). Chapter 23 Role of the locus coeruleus in emotional activation. *Prog. Brain Res.* 107, 379–402. doi:10.1016/S0079-6123(08)61877-4.

Bahrick, H. P., Bahrick, P. O., and Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *J. Exp. Psychol. Gen.* 104, 54–75. doi:10.1037/0096-3445.104.1.54.

Barnett, A. G., van der Pols, J. C., and Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *Int. J. Epidemiol.* 34, 215–220. doi:10.1093/ije/dyh299.

Von Békésy, G. (1970). Travelling Waves as Frequency Analysers in the Cochlea. *Nature* 225, 1207–1209. doi:10.1038/2251207a0.

Bermúdez-Rattoni, F. (2004). Molecular mechanisms of taste-recognition memory. *Nat. Rev. Neurosci.* 5, 209–217. doi:10.1038/nrn1344.

Besson, M., Chobert, J., and Marie, C. (2011). Transfer of Training between Music and Speech: Common Processing, Attention, and Memory. *Front. Psychol.* 2, 94. doi:10.3389/fpsyg.2011.00094.

Bichler, O., Querlioz, D., Thorpe, S. J., Bourgoin, J.-P., and Gamrat, C. (2012). Extraction of temporally correlated features from dynamic vision sensors with spike-timing-dependent plasticity. *Neural Networks* 32, 339–348. doi:10.1016/j.neunet.2012.02.022.

Blake, D. T., Hsiao, S. S., and Johnson, K. O. (1997). Neural coding mechanisms in tactile pattern recognition: the relative contributions of slowly and rapidly adapting mechanoreceptors to perceived roughness. *J. Neurosci.* 17, 7480–7489. doi:10.1109/10.81565.

Botzung, A., Denkova, E., and Manning, L. (2008). Experiencing past and future personal events: Functional neuroimaging evidence on the neural bases of mental time travel. *Brain Cogn.* 66, 202–212. doi:10.1016/j.bandc.2007.07.011.

Boyden, E. S., Katoh, A., and Raymond, J. L. (2004). CEREBELLUM-DEPENDENT LEARNING: The Role of Multiple Plasticity Mechanisms. *Annu. Rev. Neurosci.* 27, 581–609. doi:10.1146/annurev.neuro.27.070203.144238.

Boyer, P. (2008). Evolutionary economics of mental time travel? *Trends Cogn. Sci.* 12, 219–224. doi:10.1016/j.tics.2008.03.003.

Braak, H., Thal, D. R., Ghebremedhin, E., and Del Tredici, K. (2011). Stages of the Pathologic Process in Alzheimer Disease: Age Categories From 1 to 100 Years. *J. Neuropathol. Exp. Neurol.* 70.

Brett, M., Anton, J. L., Valabregue, R., and Poline, J. B. (2002). Region of interest analysis using an SPM toolbox. in *NeuroImage* (Sendai, Japan), 497. doi:http://dx.doi.org/10.1016/S1053-8119(02)90010-8.

Broadbent, D. E. (1958). "The effects of noise on behaviour.," in *Perception and communication* (Elmsford, NY, US: Pergamon Press), 81–107. Available at: http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&uid=2004-16224-005.

Brockmeier, S. J., Grasmeder, M. L., Vischer, M., Mawman, D., Baumgartner, W. D., Stark, T., et al. (2004). Comparison of musical activities by cochlear implant users. *Int. Congr. Ser.* 1273, 205–207. doi:10.1016/j.ics.2004.08.003.

Brooks, J. C. W., Faull, O. K., Pattinson, K. T. S., and Jenkinson, M. (2013). Physiological noise in brainstem FMRI. *Front. Hum. Neurosci.* 7, 623. doi:10.3389/fnhum.2013.00623.

Buchwald, J., Dickerson, L., Harrison, J., and Hinman, C. (1988). "Medial Geniculate Body Unit Responses to Cat Cries," in *Auditory Pathway* (Boston, MA: Springer US), 319–322. doi:10.1007/978-1-4684-1300-7_45.

Burianová, J., Ouda, L., and Syka, J. (2015). The influence of aging on the number of neurons and levels of non phosporylated neurofilament proteins in the central auditory system of rats. *Front. Aging Neurosci.* 7, 1–10. doi:10.3389/fnagi.2015.00027.

Bushman, B. J., and Bonacci, A. M. (2002). Violence and sex impair memory for television ads. *J. Appl. Psychol.* 87, 557–564. doi:10.1037/0021-9010.87.3.557.

Bussey, T. J., and Saksida, L. M. (2002). The organization of visual object representations: a connectionist model of effects of lesions in perirhinal cortex. *Eur. J. Neurosci.* 15, 355–364. doi:10.1046/j.0953-816x.2001.01850.x.

Bussey, T. J., Saksida, L. M., and Murray, E. A. (2006). Perirhinal cortex and feature-ambiguous discriminations. *Learn. Mem.* 13, 103-5-7. doi:10.1101/lm.163606.

Chambers, K. E., Onishi, K. H., and Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. doi:10.1016/s0010-0277(02)00233-0.

Clugnet, M. C., and LeDoux, J. E. (1990). Synaptic plasticity in fear conditioning circuits: induction of LTP in the lateral nucleus of the amygdala by stimulation of the medial geniculate body. *J. Neurosci.* 10, 2818–24. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2388089 [Accessed September 28, 2016].

Connor, C., and Johnson, K. (1992). Neural coding of tactile texture: comparison of

spatial and temporal mechanisms for roughness perception. *J. Neurosci.* 12.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia* 6, 255–265. doi:10.1016/0028-3932(68)90024-9.

Corkin, S., Amaral, D. G., Gonzá Lez, R. G., Johnson, K. A., and Hyman, B. T. (1997). H. M.'s Medial Temporal Lobe Lesion: Findings from Magnetic Resonance Imaging.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol. Bull.* 104, 163–191. doi:10.1037/0033-2909.104.2.163.

Craig, J. C., and Baihua, X. (1990). Temporal order and tactile patterns. *Percept. Psychophys.* 47, 22–34. doi:10.3758/BF03208161.

Craik, F. I. M., and Kirsner, K. (1974). The effect of speaker's voice on word recognition. *Q. J. Exp. Psychol.* 26, 274–284. doi:10.1080/14640747408400413.

Creutzfeldt, O., Hellweg, F.-C., and Schreiner, C. (1980). Thalamocortical transformation of responses to complex auditory stimuli. *Exp. Brain Res.* 39, 87–104. doi:10.1007/BF00237072.

van Dam, L. C. J., and van Ee, R. (2006). Retinal image shifts, but not eye movements per se, cause alternations in awareness during binocular rivalry. *J. Vis.* 6, 3–3. doi:10.1167/6.11.3.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.

Desmond, J. E., and Fiez, J. A. (1998a). Neuroimaging studies of the cerebellum: language, learning and memory. *Trends Cogn. Sci.* 2, 355–362. doi:10.1016/S1364-6613(98)01211-X.

Desmond, J. E., and Fiez, J. A. (1998b). Neuroimaging studies of the cerebellum: Language, learning and memory. *Trends Cogn. Sci.* 2, 355–362. doi:10.1016/S1364-6613(98)01211-X.

Desmond, J. E., Gabrieli, J. D., Wagner, A. D., Ginier, B. L., and Glover, G. H. (1997). Lobular patterns of cerebellar activation in verbal working-memory and finger-tapping tasks as revealed by functional MRI. *J. Neurosci.* 17, 9675–85. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9391022 [Accessed October 14, 2016].

Díaz, P. B. (2013). Possible biological basis for the origin of art, language, religion and science. Available at: http://www.rupestreweb.info/originofart.html.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr. Opin. Neurobiol.* 10, 732–739. doi:10.1016/S0959-4388(00)00153-7.

Doyon, J., Owen, A. M., Petrides, M., Sziklas, V., and Evans, A. C. (1996). Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography. *Eur. J. Neurosci.* 8, 637–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9081615 [Accessed October 14, 2016].

Dykes, R. W. (1997). Mechanisms controlling neuronal plasticity in somatosensory cortex. *Can. J. Physiol. Pharmacol.* 75, 535–545. doi:10.1139/y97-089.

Ebbinghaus, H., Henry, A. (Trans), and Bussenius, C. E. (Trans) (1913). *Memory: A contribution to experimental psychology*. New York: Teachers College Press Available at: http://psycnet.apa.org/books/10011/.

Edeline, J.-M., and Weinberger, N. M. (1991). Subcortical adaptive filtering in the auditory system: Associative receptive field plasticity in the dorsal medial geniculate body. *Behav. Neurosci.* 105, 154–175. doi:10.1037/0735-7044.105.1.154.

Eden, G. F., Joseph, J. E., Brown, H. E., Brown, C. P., and Zeffiro, T. A. (1999). Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: The behavior interleaved gradients technique. *Magn. Reson. Med.* 41, 13–20. doi:10.1002/(SICI)1522-2594(199901)41:1<13::AID-MRM4>3.0.CO;2-T.

Ellis, B. W., Johns, M. W., Lancaster, R., Raptopoulos, P., Angelopoulos, N., and Priest, R. G. (1981). The St. Mary's Hospital sleep questionnaire: a study of reliability. *Sleep* 4, 93–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7232974 [Accessed May 10, 2016].

Engstroem, H., Ades, H. W., and Hawkins, J. E. (1965). CELLULAR PATTERN, NERVE STRUCTURES, AND FLUID SPACES OF THE ORGAN OF CORTI. *Contrib. Sens. Physiol.* 14, 1–37. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14288211 [Accessed September 5, 2016].

Ergenzinger, E. R., Glasier, M. M., Hahm, J. O., and Pons, T. P. (1998). Cortically induced thalamic plasticity in the primate somatosensory system. *Nat. Neurosci.* 1, 226–229. doi:10.1038/673.

Evans, E. F. (1978). Place and Time Coding of Frequency in the Peripheral Auditory System: Some Physiological Pros and Cons. *Audiology*.

Frey, H.-P., Kaernbach, C., and König, P. (2003). Cats can detect repeated noise stimuli. *Neurosci. Lett.* 346, 45–48. doi:10.1016/S0304-3940(03)00559-7.

Frisina, R. D. (2001). Subcortical neural coding mechanisms for auditory temporal processing. *Hear. Res.* 158, 1–27. doi:10.1016/S0378-5955(01)00296-9.

Friston, K. J., Frith, C. D., Passingham, R. E., Liddle, P. F., and Frackowiak, R. S. J. (1992). Motor Practice and Neurophysiological Adaptation in the Cerebellum: A Positron Tomography Study. *Proc. R. Soc. London B Biol. Sci.* 248.

Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R.

S. J., et al. (1995). Analysis of fMRI Time-Series Revisited. *Neuroimage* 2, 45–53. doi:10.1006/nimg.1995.1007.

Gaffan, D. (2002). Against memory systems. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 357, 1111–21. doi:10.1098/rstb.2002.1110.

Gauthier, I., Skudlarski, P., Gore, J. C., and Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nat. Neurosci.* 3, 191–197. doi:10.1038/72140.

Gill, K. Z., and Purves, D. (2009). A Biological Rationale for Musical Scales. *PLoS One* 4, e8144. doi:10.1371/journal.pone.0008144.

Glover, G. H., Li, T. Q., and Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Med.* 44, 162–167. doi:10.1002/1522-2594(200007)44:1<162::AID-MRM23>3.0.CO;2-E.

Gold, J. M., Aizenman, A., Bond, S. M., and Sekuler, R. (2014). Memory and incidental learning for visual frozen noise sequences. *Vision Res.* 99, 19–36. doi:10.1016/j.visres.2013.09.005.

Goldberg, J. M., and Brown, P. B. (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J. Neurophysiol.* 32.

Golding, N. L., Robertson, D., and Oertel, D. (1995). Recordings from slices indicate that octopus cells of the cochlear nucleus detect coincident firing of auditory nerve fibers with temporal precision. *J. Neurosci.* 15, 3138–3153.

Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25. doi:10.1016/0166-2236(92)90344-8.

Graham, K. S., Barense, M. D., and Lee, A. C. H. (2010). Going beyond LTM in the MTL: A synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia*. doi:10.1016/j.neuropsychologia.2010.01.001 <http://dx.doi.org/10.1016/j.neuropsychologia.2010.01.001>.

Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., and Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *Nat. Neurosci.* 4, 633–637. doi:10.1038/88459.

Grothe, B., Pecka, M., and McAlpine, D. (2010). Mechanisms of Sound Localization in Mammals. *Physiol. Rev.* 90.

Guttman, N., and Julesz, B. (1963). Lower limits of auditory analysis. *J. Acoust. Soc. Am.* 35.

Hakimian, S., Anderson, C. H., and Thomas Thach, W. (1999). A PDF model of populations of Purkinje cells: Non-linear interactions and high variability.

Neurocomputing 26, 169–175. doi:10.1016/S0925-2312(99)00066-1.

Han, J.-H., Kushner, S. A., Yiu, A. P., Cole, C. J., Matynia, A., Brown, R. A., et al. (2007). Neuronal Competition and Selection During Memory Formation. *Science (80-.).* 316.

Hebb, D. (1949). *The Organization of Behavior*. New York: Wiley & Sons.

Heffner, R. S. (1997). Comparative Study of Sound Localization and its Anatomical Correlates in Mammals. *Acta Otolaryngol.* 117, 46–53. doi:10.3109/00016489709126144.

Heffner, R. S., and Heffner, H. E. (1992). Hearing and sound localization in blind mole rats (Spalax ehrenbergi). *Hear. Res.* 62, 206–216. doi:10.1016/0378-5955(92)90188-S.

Hickey, T. L., and Guillery, R. W. (1979). Variability of laminar patterns in the human lateral geniculate nucleus. *J. Comp. Neurol.* 183, 221–246. doi:10.1002/cne.901830202.

Ho, Y.-C., Cheung, M.-C., and Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology* 17, 439–450. doi:10.1037/0894-4105.17.3.439.

Hopkins, W., and Johnston, D. (1984). Frequency-dependent noradrenergic modulation of long-term potentiation in the hippocampus. *Science (80-.).* 226.

Jackson, P. L., Lafleur, M. F., Malouin, F., Richards, C. L., and Doyon, J. (2003). Functional cerebral reorganization following motor sequence learning through mental practice with motor imagery. *Neuroimage* 20, 1171–1180. doi:10.1016/S1053-8119(03)00369-0.

Jacobs, H. I. L., Wiese, S., van de Ven, V., Gronenschild, E. H. B. M., Verhey, F. R. J., Matthews, P. M., et al. (2015). Relevance of parahippocampal-locus coeruleus connectivity to memory in early dementia. *Neurobiol. Aging* 36, 618–26. doi:10.1016/j.neurobiolaging.2014.10.041.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S., and Passingham, R. E. (1994). Motor sequence learning: a study with positron emission tomography. *J. Neurosci.* 14, 3775–90. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8207487 [Accessed October 14, 2016].

Jewett, D. L., Romano, M. N., and Williston, J. S. (1970). Human Auditory Evoked Potentials: Possible Brain Stem Components Detected on the Scalp. *Science (80-.).* 167, 1517–1518. doi:10.1126/science.167.3924.1517.

Jones, B. E., and Moore, R. Y. (1977). Ascending projections of the locus coeruleus in the rat. II. Autoradiographic study. *Brain Res.* 127, 23–53. doi:10.1016/0006-8993(77)90378-X.

Jonesgotman, M., and Zatorre, R. J. (1993). Odor Recognition Memory in Humans: Role of Right Temporal and Orbitofrontal Regions. *Brain Cogn.* 22, 182–198.

doi:10.1006/brcg.1993.1033.

Jueptner, M., Frith, C. D., Brooks, D. J., Frackowiak, R. S. J., and Passingham, R. E. (1997a). Anatomy of Motor Learning. II. Subcortical Structures and Learning by Trial and Error. *J. Neurophysiol.* 77.

Jueptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S. J., and Passingham, R. E. (1997b). Anatomy of Motor Learning. I. Frontal Cortex and Attention to Action. *J. Neurophysiol.* 77.

Kaba, H., and Nakanishi, S. (1995). Synaptic Mechanisms of Olfactory Recognition Memory. *Rev. Neurosci.* 6, 125–142. doi:10.1515/REVNEURO.1995.6.2.125.

Kaernbach, C. (1992). On the consistency of tapping to repeated noise. 92, 788–793.

Kaernbach, C. (1993). Temporal and spectral basis of the features perceived in repeated noise. *J. Acoust. Soc. Am.* 94, 91–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8354764.

Kaernbach, C. (2004). The memory of noise. *Exp. Psychol.* 51, 240–8. doi:10.1027/1618-3169.51.4.240.

Kaernbach, C., and Schulze, H. (2002). Auditory sensory memory for random waveforms in the Mongolian gerbil. *Neurosci. Lett.* 329, 37–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12161257.

Kaminski, J., Call, J., and Fischer, J. (2004). Word learning in a domestic dog: evidence for "fast mapping". *Science (80-.).* 304, 1682–3. doi:10.1126/science.1097859.

Kaukoranta, E., Hari, R., and Lounasmaa, O. V. (1987). Responses of the human auditory cortex to vowel onset after fricative consonants. *Exp. Brain Res.* 69, 19–23. doi:10.1007/BF00247025.

Keane, M. M., Gabrieli, J. D. E., Mapstone, H. C., Johnson, K. A., and Corkin, S. (1995). Double dissociation of memory capacities after bilateral occipital-lobe or medial temporal-lobe lesions. *Brain* 118, 1129–1148. doi:10.1093/brain/118.5.1129.

Kim, G., Lewis-Peacock, J. a, Norman, K. a, and Turk-Browne, N. B. (2014). Pruning of memories by context-based prediction error. *Proc. Natl. Acad. Sci. U. S. A.* 111, 8997–9002. doi:10.1073/pnas.1319438111.

Kimura, A., Donishi, T., Sakoda, T., Hazama, M., and Tamai, Y. (2003). Auditory thalamic nuclei projections to the temporal cortex in the rat. *Neuroscience* 117, 1003–1016. doi:10.1016/S0306-4522(02)00949-1.

Kirkham, N. Z., Slemmer, J. A., and Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* 83, B35–B42. doi:10.1016/S0010-0277(02)00004-5.

König, P., Engel, A. K., and Singer, W. (1996). Integrator or coincidence detector? The

role of the cortical neuron revisited. *Trends Neurosci.* 19, 130–137. doi:10.1016/S0166-2236(96)80019-1.

Kraus, K. S., and Canlon, B. (2012). Neuronal connectivity and interactions between the auditory and limbic systems. Effects of noise and tinnitus. *Hear. Res.* 288, 34–46. doi:10.1016/j.heares.2012.02.009.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. 1097–1105.

Krupa, D. J., Ghazanfar, A. A., and Nicolelis, M. A. (1999). Immediate thalamic sensory plasticity depends on corticothalamic feedback. *Proc. Natl. Acad. Sci. U. S. A.* 96, 8200–5. doi:10.1073/PNAS.96.14.8200.

Kumar, S., Bonnici, H. M., Teki, S., Agus, T. R., Pressnitzer, D., Maguire, E. A., et al. (2014). Representations of specific acoustic patterns in the auditory cortex and hippocampus. *Proc. Biol. Sci.* 281, 20141000. doi:10.1098/rspb.2014.1000.

Lachter, J., Forster, K. I., and Ruthruff, E. (2004). Forty-five years after Broadbent (1958): still no identification without attention. *Psychol. Rev.* 111, 880–913. doi:10.1037/0033-295X.111.4.880.

Lafleur, M. F., Jackson, P. L., Malouin, F., Richards, C. L., Evans, A. C., and Doyon, J. (2002). Motor Learning Produces Parallel Dynamic Functional Changes during the Execution and Imagination of Sequential Foot Movements. *Neuroimage* 16, 142–157. doi:10.1006/nimg.2001.1048.

Lamprecht, R., and LeDoux, J. (2004). Structural plasticity and memory. *Nat. Rev. Neurosci.* 5, 45–54. doi:10.1038/nrn1301.

Lavezzi, A. M., Pusiol, T., and Matturri, L. (2015). Cytoarchitectural and functional abnormalities of the inferior colliculus in sudden unexplained perinatal death. *Medicine (Baltimore).* 94, e487. doi:10.1097/MD.0000000000000487.

Law, L. N. C., Zentner, M., Thomson, J., Goswami, U., Helmbold, N., Rammsayer, T., et al. (2012). Assessing Musical Abilities Objectively: Construction and Validation of the Profile of Music Perception Skills. *PLoS One* 7, e52508. doi:10.1371/journal.pone.0052508.

Lederman, S. J., and Klatzky, R. L. (2009). Haptic perception: A tutorial. *Atten. Percept. Psychophys.* 71, 1439–1459. doi:10.3758/APP.71.7.1439.

Limbert, C., and Patterson, R. D. (1982). Tapping to repeated noise. *J. Acoust. Soc. Am.* 71, S38. doi:10.1121/1.2019365.

Lømo, T. (1971a). Patterns of activation in a monosynaptic cortical pathway: The perforant path input to the dentate area of the hippocampal formation. *Exp. Brain Res.* 12, 18–45. doi:10.1007/BF00234414.

Lømo, T. (1971b). Potentiation of monosynaptic EPSPs in the perforant path-dentate

granule cell synapse. *Exp. Brain Res.* 12, 46–63. doi:10.1007/BF00234415.

Lømo, T. (2003). The discovery of long-term potentiation. *Philos. Trans. R. Soc. London B Biol. Sci.* 358.

Luo, H., and Poeppel, D. (2012). Cortical Oscillations in Auditory Perception and Speech: Evidence for Two Temporal Windows in Human Auditory Cortex. *Front. Psychol.* 3, 170. doi:10.3389/fpsyg.2012.00170.

Luo, H., Tian, X., Song, K., Zhou, K., and Poeppel, D. (2013). Neural response phase tracks how listeners learn new acoustic representations. *Curr. Biol.* 23, 968–974. doi:10.1016/j.cub.2013.04.031.

Manjarrez, E., Rojas-Piloni, G., Méndez, I., and Flores, A. (2003). Stochastic resonance within the somatosensory system: effects of noise on evoked field potentials elicited by tactile stimuli. *J. Neurosci.* 23, 1997–2001.

Marshall, L., Helgadóttir, H., Mölle, M., and Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature* 444, 610–613. doi:10.1038/nature05278.

Mary, A., Schreiner, S., and Peigneux, P. (2013). Accelerated long-term forgetting in aging and intra-sleep awakenings. *Front. Psychol.* 4, 750. doi:10.3389/fpsyg.2013.00750.

Masquelier, T., Guyonneau, R., and Thorpe, S. J. (2008). Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains. *PLoS One* 3, e1377. doi:10.1371/journal.pone.0001377.

Masquelier, T., Guyonneau, R., and Thorpe, S. J. (2009). Competitive STDP-based spike pattern learning. *Neural Comput.* 21, 1259–76. doi:10.1162/neco.2008.06-08-804.

Masquelier, T., Umr, C., and Toulouse, C. U. (2016). STDP allows close-to-optimal multi-neuron spike sequence detection by single coincidence detector neurons. 323711, 323711.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*

Moerel, M., De Martino, F., Uğurbil, K., Yacoub, E., and Formisano, E. (2015). Processing of frequency and location in human subcortical auditory structures. *Sci. Rep.* 5, 17048. doi:10.1038/srep17048.

Moore, J. K. (1987). The human auditory brain stem: A comparative view. *Hear. Res.* 29, 1–32. doi:10.1016/0378-5955(87)90202-4.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Q. J. Exp. Psychol.* 11, 56–60. doi:10.1080/17470215908416289.

Moray, N., and O'Brien, T. (1967). Signal-Detection Theory Applied to Selective Listening. *J. Acoust. Soc. Am.* 42, 765. doi:10.1121/1.1910648.

Moser, T., Neef, A., and Khimich, D. (2006). Mechanisms underlying the temporal precision of sound coding at the inner hair cell ribbon synapse. *J. Physiol.* 576, 55–62. doi:10.1113/jphysiol.2006.114835.

Murre, J. M. J., Dros, J., Anderson, J., Schooler, L., Rubin, D., Wenzel, A., et al. (2015). Replication and Analysis of Ebbinghaus' Forgetting Curve. *PLoS One* 10, e0120644. doi:10.1371/journal.pone.0120644.

Nadel, L., Samsonovich, A., Ryan, L., and Moscovitch, M. (2000). Multiple trace theory of human memory: Computational, neuroimaging, and neuropsychological results. *Hippocampus* 10, 352–368. doi:10.1002/1098-1063(2000)10:4<352::AID-HIPO2>3.0.CO;2-D.

Nader, K., Schafe, G. E., and Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidationafter retrieval. *Nature* 406, 722–726. doi:10.1038/35021052.

Nelken, I., and Young, E. D. (1996). Why do cats need a dorsal cochlear nucleus? *J. Basic Clin. Physiol. Pharmacol.* 7, 199–220. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8910137 [Accessed October 7, 2016].

Nielsen, F. Å., Balslev, D., and Hansen, L. K. (2005). Mining the posterior cingulate: Segregation between memory and pain components. *Neuroimage* 27, 520–532. doi:10.1016/j.neuroimage.2005.04.034.

O'Doherty, J., Rolls, E. T., Francis, S., Bowtell, R., and McGlone, F. (2001). Representation of Pleasant and Aversive Taste in the Human Brain. *J. Neurophysiol.* 85.

Ode, H., Nakashima, M., Kitamura, S., Sugiura, W., and Sato, H. (2012). Molecular dynamics simulation in virus research. *Front. Microbiol.* 3, 258. doi:10.3389/fmicb.2012.00258.

Olsen, J. F. (1994). Medial geniculate neurons in the squirrel monkey sensitive to inter-component delays that categorize species-typical calls. in *Assoc Res Otolaryngol Abst*, Abst 17:21.

Olshausen, B. A., and Field, D. J. (2004). Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.* 14, 481–487. doi:10.1016/j.conb.2004.07.007.

Olszewski, J., and Baxter, D. (2014). *Olszewski and Baxter's Cytoarchitecture of the Human Brainstem*. 3rd, revis ed. , eds. Büttner-Ennever J.A. (Munich) and Horn A.K.E. (Munich) Karger Publishers Available at: http://www.karger.com/Book/Home/259866.

Pantev, C., and Herholz, S. C. (2011). Plasticity of the human auditory cortex related to musical training. *Neurosci. Biobehav. Rev.* 35, 2140–2154. doi:10.1016/j.neubiorev.2011.06.010.

Papez, J. W. (1937). A proposed mechanism of emotion. *Arch. Neurol. Psychiatry* 38, 725–743. doi:10.1176/jnp.7.1.103.

Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., and Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36, 767–76. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12441063.

Payne, J. D. (2011). Sleep on it!: stabilizing and transforming memories during sleep. *Nat. Neurosci.* 14, 272–274. doi:10.1038/nn0311-272.

Pickel, V. M., Segal, M., and Bloom, F. E. (1974). A radioautographic study of the efferent pathways of the nucleus locus coeruleus. *J. Comp. Neurol.* 155, 15–41. doi:10.1002/cne.901550103.

Pilley, J. W., and Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behav. Processes* 86, 184–195. doi:10.1016/j.beproc.2010.11.007.

Pollack, I. (1972). Memory for auditory waveform. *J. Acoust. Soc. Am.* 52, 1209–15. Available at: http://www.ncbi.nlm.nih.gov/pubmed/4638035.

Pollack, I., and Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychon. Sci.* 1, 125–126.

Pontes, C., Reis, F. F., and Sousa-Pinto, A. (1975). The auditory cortical projections onto the medial geniculate body in the cat. An experimental anatomical study with silver and autoradiographic methods. *Brain Res.* 91, 43–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1131700 [Accessed October 18, 2016].

Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., McNamara, J. O., et al. (2008). *Neuroscience*. 4th ed. Sinauer Associates.

Quian Quiroga, R., Kraskov, A., Koch, C., and Fried, I. (2009). Explicit Encoding of Multimodal Percepts by Single Neurons in the Human Brain. doi:10.1016/j.cub.2009.06.060.

Quiroga, R. Q., Fried, I., and Koch, C. (2013). Brain Cells for Grandmother. *Sci. Am.* 308, 30–35. doi:10.1038/scientificamerican0213-30.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature* 435, 1102–1107. doi:10.1038/nature03687.

Rajendran, V. G., Harper, N. S., Abdel-Latif, K. H. A., and Schnupp, J. W. H. (2016). Rhythm facilitates the detection of repeating sound patterns. *Front. Neurosci.* 10, 1–7. doi:10.3389/fnins.2016.00009.

Rando, O. J., Verstrepen, K. J., Acar, M., Becskei, A., Oudenaarden, A. van, Aertsen, A., et al. (2007). Timescales of genetic and epigenetic inheritance. *Cell* 128, 655–68. doi:10.1016/j.cell.2007.01.023.

Raphael, Y., and Altschuler, R. A. (2003). Structure and innervation of the cochlea. *Brain Res. Bull.* 60, 397–422. doi:10.1016/S0361-9230(03)00047-9.

Rawlins, J. N. P., Aggleton, J. P., Mishkin, M., Alexander, M. P., Freedman, M., Amsel, A., et al. (1985). Associations across time: The hippocampus as a temporary memory store. *Behav. Brain Sci.* 8, 479. doi:10.1017/S0140525X00001291.

Reite, M., Adams, M., Simon, J., Teale, P., Sheeder, J., Richardson, D., et al. (1994). Auditory M100 component 1: relationship to Heschl's gyri. *Brain Res. Cogn. Brain Res.* 2, 13–20. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7812174 [Accessed October 14, 2016].

Rolls, E. T. (2013). The mechanisms for pattern completion and pattern separation in the hippocampus. *Front. Syst. Neurosci.* 7, 74. doi:10.3389/fnsys.2013.00074.

Rolls, E. T., and Treves, A. (1990). The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain. *Network*.

Rubin, D. C., and Kontis, T. C. (1983). A schema for common cents. *Mem. Cognit.* 11, 335–341. doi:10.3758/BF03202446.

Rudoy, J. D., Voss, J. L., Westerberg, C. E., and Paller, K. A. (2009). Strengthening Individual Memories by Reactivating them During Sleep. *Science (80-.).* 29, 997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted.

Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., and Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behav. Brain Res.* 156, 95–103. doi:10.1016/j.bbr.2004.05.012.

Ryan, L., Nadel, L., Keil, K., Putnam, K., Schnyer, D., Trouard, T., et al. (2001). Hippocampal complex and retrieval of recent and very remote autobiographical memories: Evidence from functional magnetic resonance imaging in neurologically intact people. *Hippocampus* 11, 707–714. doi:10.1002/hipo.1086.

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science (80-.).* 274.

Särkkä, S., Solin, A., Nummenmaa, A., Vehtari, A., Auranen, T., Vanni, S., et al. (2012). Dynamic retrospective filtering of physiological noise in BOLD fMRI: DRIFTER. *Neuroimage* 60, 1517–27. doi:10.1016/j.neuroimage.2012.01.067.

Schönwiesner, M., Novitski, N., Pakarinen, S., Carlson, S., Tervaniemi, M., and Näätänen, R. (2007). Heschl's gyrus, posterior superior temporal gyrus, and mid-ventrolateral prefrontal cortex have different roles in the detection of acoustic changes. *J. Neurophysiol.* 97, 2075–82. doi:10.1152/jn.01083.2006.

Scoviille, W. B., and Milner, B. (1957). LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. *J. Neurol. Neurosurg. Psychiat* 20.

Sejnowski, T. J., and Destexhe, A. (2000). Why do we sleep? *Brain Res.* 886, 208–223.

doi:10.1016/S0006-8993(00)03007-9.

Shibata, H., and Yukie, M. (2003). Differential thalamic connections of the posteroventral and dorsal posterior cingulate gyrus in the monkey. *Eur. J. Neurosci.* 18, 1615–1626. doi:10.1046/j.1460-9568.2003.02868.x.

Skoe, E., and Kraus, N. (2010). Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* 31, 302–24. doi:10.1097/AUD.0b013e3181cdb272.

Skosnik, P. D., Mirza, F., Gitelman, D. R., Parrish, T. B., Mesulam, M.-M., and Reber, P. J. (2002). Neural Correlates of Artificial Grammar Learning. *Neuroimage* 17, 1306–1314. doi:10.1006/nimg.2002.1291.

Smith, E. C., and Lewicki, M. S. (2006). Efficient auditory coding. *Nature* 439, 978–982. doi:10.1038/nature04485.

Song, J. H., Skoe, E., Wong, P. C. M., and Kraus, N. (2008). Plasticity in the Adult Human Auditory Brainstem following Short-term Linguistic Training. *J. Cogn. Neurosci.* 20, 1892–1902. doi:10.1162/jocn.2008.20131.

Squire, L. R., Amaral, D. G., Zola-Morgan, S., Kritchevsky, M., and Press, G. (1989). Description of brain injury in the amnesic patient N.A. Based on magnetic resonance imaging. *Exp. Neurol.* 105, 23–35. doi:10.1016/0014-4886(89)90168-4.

Squire, L. R., Cohen, N. J., and Nadel, L. (1984). "The medial temporal region and memory consolidation: A new hypothesis," in *Memory consolidation: Psychobiology of Cognition*, eds. H. Weingartner and E. S. Parker Available at: https://books.google.fr/books?hl=en&lr=&id=pfohAwAAQBAJ&oi=fnd&pg=PA185&dq=squire+AND+1984&ots=jkFY2yqAkV&sig=Hbv9-Hl62X_ZqqvW3LLutnb69NA#v=onepage&q=squire AND 1984&f=false.

Stanislaw, H., and Todorov, N. (1999). Calculation of signal detection theory measures. *Behav. Res. Methods. Instrum. Comput.* 31, 137–49. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10495845 [Accessed January 23, 2014].

Stanton, P. K., and Sarvey, J. M. (1987). Norepinephrine regulates long-term potentiation of both the population spike and dendritic EPSP in hippocampal dentate gyrus. *Brain Res. Bull.* 18, 115–119. doi:10.1016/0361-9230(87)90039-6.

Stanton, P., and Sarvey, J. (1985). Depletion of norepinephrine, but not serotonin, reduces long-term potentiation in the dentate gyrus of rat hippocampal slices. *J. Neurosci.* 5.

Steriade, M., Jones, E. G., and Llinás, R. R. (1990). *Thalamic oscillations and signaling*. Wiley, New York.

Steriade, M., McCormick, D. A., and Sejnowski, T. J. (1993). Thalamocortical Oscillations in the Sleeping and Aroused Brain.

Strait, D. L., and Kraus, N. (2014). Biological impact of auditory expertise across the life

span: musicians as a model of auditory learning. *Hear. Res.* 308, 109–21. doi:10.1016/j.heares.2013.08.004.

Strange, B. A., Fletcher, P. C., Henson, R. N., Friston, K. J., and Dolan, R. J. (1999). Segregating the functions of human hippocampus. *Proc. Natl. Acad. Sci. U. S. A.* 96, 4034–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10097158 [Accessed October 13, 2016].

Suddendorf, T., and Busby, J. (2003). Mental time travel in animals? *Trends Cogn. Sci.* 7, 391–396. doi:10.1016/S1364-6613(03)00187-6.

Suddendorf, T., and Corballis, M. C. (1997). Mental time travel and the evolution of the human mind.

Suga, N., Gao, E., Zhang, Y., Ma, X., and Olsen, J. F. (2000). The corticofugal system for hearing: recent progress. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11807–14. doi:10.1073/pnas.97.22.11807.

Suga, N., Yan, J., and Zhang, Y. (1998). "The Processing of Species-Specific Complex Sounds by the Ascending and Descending Auditory Systems," in *Central Auditory Processing and Neural Modeling* (Boston, MA: Springer US), 55–70. doi:10.1007/978-1-4615-5351-9_6.

Sweatt, D. J. (2010). *Mechanisms of Memory*. 2nd ed. Academic press, Elsevier Available at: https://books.google.fr/books?hl=en&lr=&id=q4GGGCzgGF8C&oi=fnd&pg=PP1&dq=mechanisms+of+memory+sweatt&ots=Dtt5o4dnQ4&sig=rEZt46EP0VRaYnP9z2M3U-BKu64#v=onepage&q=mechanisms of memory sweatt&f=false.

Syka, J., Popelář, J., Kvašňák, E., and Šuta, D. (1998). "Processing of Vocalization Signals in Neurons of the Inferior Colliculus and Medial Geniculate Body," in *Central Auditory Processing and Neural Modeling* (Boston, MA: Springer US), 1–11. doi:10.1007/978-1-4615-5351-9_1.

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., et al. (2006). Declarative memory consolidation in humans : A prospective functional magnetic resonance. 2–7.

Tanaka, H., and Taniguchi, I. (1991). Responses of Medial Geniculate Neurons to Species-Specific Vocalized Sounds in the Guinea Pig. *Jpn. J. Physiol.* 41, 817–829. doi:10.2170/jjphysiol.41.817.

Teuber, H.-L., Milner, B., and Vaughan, H. G. (1968). Persistent anterograde amnesia after stab wound of the basal brain. *Neuropsychologia* 6, 267–282. doi:10.1016/0028-3932(68)90025-0.

Thorpe, S. J., Rolls, E. T., and Maddison, S. (1983). The orbitofrontal cortex: Neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115. doi:10.1007/BF00235545.

Tinbergen, and Niko (1953). The herring gull's world: a study of the social behaviour of birds.

Treisman, A., and Geffen, G. (1967). Selective attention: Perception or response? *Q. J. Exp. Psychol.* 19, 1–17. doi:10.1080/14640746708400062.

Tsoukatos, J., Kiss, Z. H. T., Davis, K. D., Tasker, R. R., and Dostrovsky, J. O. (1997). Patterns of neuronal firing in the human lateral thalamus during sleep and wakefulness. *Exp. Brain Res.* 113, 273–282. doi:10.1007/BF02450325.

Turk-Browne, N. B., Yi, D.-J., and Chun, M. M. (2006). Linking implicit and explicit memory: common encoding factors and shared representations. *Neuron* 49, 917–27. doi:10.1016/j.neuron.2006.01.030.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. doi:10.1006/nimg.2001.0978.

Udell, M. A. R., Dorey, N. R., and Wynne, C. D. L. (2008). Wolves outperform dogs in following human social cues. *Anim. Behav.* 76, 1767–1773. doi:10.1016/j.anbehav.2008.07.028.

VanRullen, R., and Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn. Sci.* 7, 207–213. doi:10.1016/S1364-6613(03)00095-0.

VanRullen, R., Zoefel, B., and Ilhan, B. (2014). On the cyclic nature of perception in vision versus audition. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369, 20130214. doi:10.1098/rstb.2013.0214.

Villa, A. E. P., Rouiller, E. M., Simm, G. M., Zurita, P., de Ribaupierre, Y., and de Ribaupierre, F. (1991). Corticofugal modulation of the information processing in the auditory thalamus of the cat. *Exp. Brain Res.* 86, 506–517. doi:10.1007/BF00230524.

Viswanathan, J., Rémy, F., Bacon-Macé, N., and Thorpe, S. J. (2016). Long term memory for noise: evidence of robust encoding of very short temporal acoustic patterns. *Front. Neurosci.* 10, 490. doi:10.3389/FNINS.2016.00490.

Vogt, B. A., Vogt, L., and Laureys, S. (2006). Cytology and functionally correlated circuits of human posterior cingulate areas. *Neuroimage* 29, 452–66. doi:10.1016/j.neuroimage.2005.07.048.

Wagner, A. D., Shannon, B. J., Kahn, I., and Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* 9, 445–453. doi:10.1016/j.tics.2005.07.001.

Walker, M., and Stickgold, R. (2014). Sleep, memory and plasticity. *Neurosci. Psychoanal.* Available at: https://books.google.fr/books?hl=en&lr=&id=SBlGBAAAQBAJ&oi=fnd&pg=PA9

3&dq=role+of+sleep+on+memory&ots=gTJY2mxMBU&sig=PMitqAhPEMDOjEsCmIgN8c6CpcM [Accessed December 10, 2015].

Warren, R. M., Bashford, J. A., Cooley, J. M., and Brubaker, B. S. (2001). Detection of acoustic repetition for very long stochastic patterns. *Percept. Psychophys.* 63, 175–182.

Wiesenfeld, K., and Moss, F. (1995). Stochastic Resonance and the Benefits of Noise. *Nature* 373.

Willander, J., and Baraldi, S. (2010). Development of a new Clarity of Auditory Imagery Scale. *Behav. Res. Methods* 42, 785–90. doi:10.3758/BRM.42.3.785.

Winer, J. A. (1984). The human medial geniculate body. *Hear. Res.* 15, 225–247. doi:10.1016/0378-5955(84)90031-5.

Winkler, I., and Cowan, N. (2005). From sensory to long-term memory: Evidence from auditory memory reactivation studies. *Exp. Psychol.* 52, 3–20. doi:10.1027/1618-3169.52.1.3.

Winkler, I., Korzyukov, O., Gumenyuk, V., Cowan, N., Linkenkaer-Hansen, K., Ilmoniemi, R. J., et al. (2002). Temporary and longer term retention of acoustic information. *Psychophysiology* 39, S0048577201393186. doi:10.1017/S0048577201393186.

Yan, J., and Suga, N. (1996). Corticofugal Modulation of Time-Domain Processing of Biosonar Information in Bats. *Science (80-.).* 273.

Yan, J., and Suga, N. (1999). Corticofugal Amplification of Facilitative Auditory Responses of Subcortical Combination-Sensitive Neurons in the Mustached Bat. *J. Neurophysiol.* 81.

Yan, W., and Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditorysystem. *Nat. Neurosci.* 1, 54–58. doi:10.1038/255.

Yassa, M. A., and Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends Neurosci.* 34, 515–25. doi:10.1016/j.tins.2011.06.006.

Yiu, A. P., Mercaldo, V., Yan, C., Richards, B., Rashid, A. J., Hsiang, H.-L. L., et al. (2014). Neurons Are Recruited to a Memory Trace Based on Relative Neuronal Excitability Immediately before Training. *Neuron* 83, 722–735. doi:10.1016/j.neuron.2014.07.017.

Young, E. D., and Nelken, I. (1998). "Interneurons Which Shape Response Properties in Dorsal Cochlear Nucleus," in *Central Auditory Processing and Neural Modeling* (Boston, MA: Springer US), 101–115. doi:10.1007/978-1-4615-5351-9_10.

Young, E. D., and Oertel, D. (2003). "The Cochlear Nucleus.," in *The synaptic organization of the brain*, ed. G. M. Shepherd (Oxford University Press), 125–163.

Yukie, M. (1995). Neural connections of auditory association cortex with the posterior cingulate cortex in the monkey. *Neurosci. Res.* 22, 179–187. doi:10.1016/0168-0102(95)00888-1.

Zatorre, R. J., Evans, A. C., Meyer, E., and Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–9. doi:10.1126/science.256.5058.846.

Zhang, Y., and Suga, N. (2000). Modulation of Responses and Frequency Tuning of Thalamic and Collicular Neurons by Cortical Activation in Mustached Bats. *J. Neurophysiol.* 84.

Zhang, Y., Suga, N., and Yan, J. (1997). Corticofugal modulation of frequency processing in bat auditory system. *Nature* 387, 900–3. doi:10.1038/43180.

Zhou, Y.-D., and Fuster, J. M. (1996). Mnemonic neuronal activity in somatosensory cortex. *Psychology* 93, 10533–10537.

Zoefel, B., Reddy Pasham, N., Brüers, S., and VanRullen, R. (2015). The ability of the auditory system to cope with temporal subsampling depends on the hierarchical level of processing. *Neuroreport* 26, 773–8. doi:10.1097/WNR.0000000000000422.

Zouridakis, G., Simos, P. G., and Papanicolaou, A. C. (1998). Multiple bilaterally asymmetric cortical sources account for the auditory N1m component. *Brain Topogr.* 10, 183–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9562539 [Accessed October 13, 2016].

Zucker, R. S., Delaney, K. R., Mulkey, R., Tank, D. W., and Poo, M. (1991). Presynaptic calcium in transmitter release and posttetanic potentiation. *Ann. N. Y. Acad. Sci.* 635, 191–207. doi:10.1038/25665.

VII. ACRONYMS

GENERAL

Convolutional Neural Network ..... CNN

Deoxyribonucleic Acid ..... DNA

Gamma Amino Butyric Acid ..... GABA

Inter-aural Level Differences ..... ILD

Inter-aural Time Differences ..... ITD

Long-term memory ..... LTM

Long term potentiation ..... LTP

Rapid Eye Movement Sleep ..... REM

Short-term memory ..... STM

Slow Wave Sleep ..... SWS

TASK RELATED

Auditory Brainstem Response ..... ABR

Cyclic Noise ..... CN

Electroencephalography ..... EEG

Event related potential ..... ERP

Event related spectral perturbation ..... ERSP

Functional Magnetic Resonance Imaging ..... fMRI

Inter Trial phase Coherence ..... ITC

Noise (plain/non-cyclic) ..... N

NEUROANATOMIC

Cochlear Nucleus ..... CoN

Heschl's Gyrus ..... HG

Hippocampus ..... HC

Inferior Colliculus IC

Locus Coeruleus LC

Medial Geniculate Body ..... MGB

Medial Temporal Lobe ..... MTL

Posterior cingulate cortex ..... PCC

Superior Olivary Complex ..... SOC

COMPUTATIONAL MODELLING

Leaky Integrate and Fire (neuron) ..... LIF

Spike-time dependent plasticity ..... STDP

VIII. RÉSUMÉ SUBSTANTIEL

Depuis l'aube de la civilisation, l'importance de la mémoire dans la pensée, la créativité, le savoir et la prise de décision a été reconnue et des philosophes de renom ainsi que des savants ont mis en avant des théories sur le possible fonctionnement de la mémoire. Malgré l'histoire riche de cette quête philosophique, ce n'est qu'en 1879 qu'Ebbinghaus initia l'étude systématique des mécanismes de la mémoire par l'utilisation de paradigmes expérimentaux. Par la suite, des décennies de recherche ont été menées, cherchant à déterminer le fonctionnement de la mémoire, et ont abouti à la définition des mécanismes d'acquisition, de traitement, de stockage (encodage), de stabilisation (consolidation) et de rappel de l'information qui ont été classifiés de différentes manières. Une classification de la mémoire a été réalisée, basée sur la durée de rétention de l'information dans le cerveau (mémoire immédiate, de travail, et à long terme) en se basant sur l'accès à l'information par des processus conscients (mémoire déclarative et non déclarative) ainsi que sur l'échelle de temps des événements (mémoire moléculaire, épigénétique et évolutive). Il est facile de postuler que ces distinctions sont rigides, mais lorsque l'on considère la mémoire en termes de voyage mental dans le temps il est clair que cela n'est pas le cas. Le voyage mental dans le temps, ou la capacité à utiliser des informations précédemment encodées afin de créer des prédictions sur de futurs événements en se basant sur des probabilités, est un phénomène observé chez l'homme et l'animal. La recherche dans ce domaine a démontré que les zones du cerveau impliquées dans la rétrospection et la prospection sont identiques, suggérant ainsi que les mécanismes de mémoire sont souples et que les aires recrutées dans une fonction sont basées sur les calculs opérés par les neurones contenus au sein de ces régions. Cette idée a été renforcée par le travail réalisé sur l'apprentissage des habiletés motrices qui montre qu'avec l'apprentissage, au fur et à mesure que les mouvements deviennent faciles à exécuter, la mémoire "se déplace" dans le cerveau avec le recrutement d'aires cérébrales impliquées dans des séquences plus automatiques. De façon similaire, alors qu'il est facile de postuler que la perception et la mémoire sont différentes et ont été traditionnellement étudiées comme des phénomènes distincts, des études récentes suggèrent que la perception sensorielle et la mémoire sont intrinsèquement liées. Les études de neuroimagerie et de comportement chez des patients atteints de lésions cérébrales (dans différentes aires) ont montré un lien étroit entre des dysfonctionnements dans le traitement sensoriel et une perte de mémoire sensorielle, ce qui a mené à une vue non modulaire de la mémoire et de la perception. La compréhension des processus de traitement sensoriel – ou comment l'information sensorielle est extraite de l'entrée sensorielle – est par conséquent un aspect important à considérer afin de comprendre comment cette information est par la suite stockée à court et long terme.

Les mécanismes de la mémoire peuvent être étudiés à différents niveaux d'organisation : la compréhension des évènements de chaque niveau d'organisation aidant ainsi à comprendre comment les souvenirs fonctionnent dans leur ensemble, tel un puzzle dont les pièces seraient réunies grâce à l'utilisation de résultats de différents

domaines. Au niveau cellulaire, la plasticité synaptique [médiée par la potentialisation à long terme (LTP ou Long Term Potentiation) et la plasticité fonction de l'occurrence des potentiels d'action ou STDP (Spike Time Dependent Plasticity)] a été identifiée comme le mécanisme principal du stockage de l'information. La mesure des mécanismes de la LTP dans l'hippocampe et le lobe temporal médian ont permis l'identification de régions clés dans les mécanismes d'encodage et de consolidation de la mémoire déclarative, un résultat qui a depuis été confirmé par des études menées chez des patients comme HM et NA possédant des lésions dans ces aires cérébrales. La synthèse de ces résultats a permis d'émettre l'hypothèse que l'hippocampe est impliqué dans le "patron de séparation" et le "patron de réalisation" de ces opérations, via la comparaison des données nouvellement acquises avec les données précédemment collectées. Par conséquent, comme prédit par le modèle représentationnel hiérarchique, ou « emergent memory account », l'hippocampe est recruté à la fois pour le traitement perceptuel et mnésique en fonction des calculs requis par la tâche. Cette notion est en accord avec l'idée que l'origine du développement des divers mécanismes de mémoire repose sur une augmentation des prérequis computationnels des tâches au cours de l'évolution.

La réponse à la question "comment l'information est contenue dans un réseau de neurones ?" repose sur la compréhension des mécanismes de codage épars et distribués. Les mécanismes de codage distribués postulent que l'information est réduite et contenue dans un réseau de neurones alors que le mécanisme de codage épars repose sur le fait que l'idée / concept à stocker est contenu au sein d'un nombre restreint de neurones qui se sont spécialisés pour répondre à ce stimulus. La mise en évidence du neurone "Jennifer Aniston" de même que les expériences de biologie moléculaire analysant les marqueurs spécifiques de l'activité neuronale ont renforcé l'idée d'un mécanisme de codage épars.

Deux paramètres, sommeil et attention sont connus pour influencer l'encodage et la consolidation de l'information et sont à considérer dans l'investigation des mécanismes de mémoire. Trois méthodologies expérimentales principales ont été utilisées pour tester les hypothèses posées sur les mécanismes de la mémoire: (i) Les expériences étudiant la mémoire pour des stimuli qui ont été encodés naturellement, avant toute expérimentation, (ii) les expériences étudiant la mémoire pour des stimuli qui ont été encodés puis testés après une période déterminée dans le laboratoire ou (iii) les expériences étudiant la mémoire des patients ayant des lésions cérébrales utilisant des stimuli réels ou artificiels. Bien que chacune de ces méthodologies possède des avantages et des inconvénients, au cours de ce travail de thèse nous avons voulu comprendre, chez des participants sains, les mécanismes de la mémoire pour des stimuli auditifs sans signification. Les stimuli ont été encodés et restitués dans le contexte du laboratoire.

Déterminer les limites de la perception et le début de la mémoire n'est pas chose facile tant d'un point de vue philosophique que neurobiologique. En effet, la compréhension de ce qui est perçu requiert des "échantillons" ou souvenirs, et les souvenirs constituent des ensembles de perceptions sensorielles. La capacité à stocker des informations

sensorielles pures a été démontrée dans toutes les modalités sensorielles mais les mécanismes par lesquels cette information est stockée restent mal définis. Pendant longtemps, il a été admis que la mémoire sensorielle se formait indépendamment de l'attention, possédait une résolution fine et un temps de rétention court et était spécifique de la modalité sensorielle. Toutefois, de nombreuses évidences ont mis à mal cette hypothèse et ont suscité un besoin pour un nouveau modèle pouvant expliquer le stockage à long-terme d'une information purement sensorielle.

En figeant un segment de bruit blanc et en le jouant de façon répétée, Guttman and Julesz (Guttman, N., and Julesz, 1963) ont créé un bruit cyclique (CN, cyclic noise). Ils ont ainsi pu montrer qu'après quelques présentations, les participants avaient la capacité de percevoir une cyclicité en détectant une information comme paramètre se répétant de façon rythmique. Les participants identifièrent ces caractéristiques, ou brefs percepts auditifs comme des « cliquetis » et des « sifflements ». Le paradigme « frozen noise » (Guttman, N., and Julesz, 1963) a été à l'origine des stimuli utilisés dans les expériences décrites dans cette thèse, qui s'est attachée à étudier la mémoire sensorielle dans la modalité auditive. Ces stimuli comblent le vide existant entre les systèmes de mémoire déclarative et non déclarative, ne possèdent aucune fluctuation identifiable en amplitude qui pourrait rendre un segment marquant comparé aux autres et forment les blocks d'apprentissage observés pendant la petite enfance lorsque tous les stimuli sont sans signification au premier abord.

La capacité des systèmes artificiels comme les réseaux profonds convolutionnels d'exécuter des taches à l'échelle de l'homme voir à des meilleurs niveaux, résout un problème d'un point de vue computationnel. Cependant il reste à déterminer si des mécanismes computationnels similaires sont utilisés par les systèmes biologiques lors la réalisation de tâches similaires. Un des mécanismes fondamentaux par lequel l'apprentissage non supervisé est réalisé dans les systèmes biologiques est la plasticité fonction d'occurrence des impulsions ou STDP, une forme de plasticité synaptique Hebbienne. Plus récemment, STDP a inspiré les neuroscientifiques computationnels et les ingénieurs à développer des solutions biologiquement inspirées pour des tâches computationnelles. Récemment, les modèles STDP ont été utilisés pour l'apprentissage de segments répétés de bruit continu (Masquelier et al., 2008, 2009, 2016). En implémentant un neurone LIF (Leaky Integrate-and-Fire) capable de détecter les patterns de potentiels d'action dans un flux d'entrée continu de bruit contenant un motif arbitraire répété aléatoirement, les auteurs ont démontré que le neurone achevait 100% de sélectivité pour répondre à la cible (motif arbitraire) en moins de 10 présentations, démontrant une capacité d'apprentissage rapide et non supervisée. Ainsi, STDP pourrait être le mécanisme candidat par lequel les neurones emmagasinent de l'information sans signification, simplement par répétition de cette information.

Sur la base des idées précédemment citées, comme l'encodage épars de la mémoire, la spécialisation des neurones dans la détection de certains traits, et la STDP comme probable mécanisme par lequel les neurones réalisent l'apprentissage des patrons de sons répétés, un de mes directeurs de thèse, Dr Simon Thorpe, a développé une étude visant à tester 10 postulats expérimentaux provocateurs en lien avec les mécanismes de

mémorisation. Ce projet de recherche a été financé par le Conseil Européen de la recherche en tant que projet M4 – Memory Mechanisms in Man and Machine (Mécanismes de la mémoire chez l'homme et la machine). Les expériences menées au cours de mes trois années de thèse ont eu pour objectif de tester explicitement certains de ces postulats : - la reconnaissance de stimuli encodés est possible sans réactiver la trace mnésique dans l'intervalle de temps, - pendant la mémorisation, la force de l'encodage augmente de façon linéaire avec le nombre de présentations, - l'attention peut aider au stockage de la mémoire de façon efficace, - la mémoire peut perdurer toute la vie, - et le modèle STDP peut expliquer ce mécanisme.

Le système auditif réalise une prouesse en transformant une énergie mécanique en vagues de sons qui deviennent percepts. La précision temporelle est essentielle dans l'audition afin de localiser l'origine du son dans l'espace. En provenance directe de la cochlée, organe auditif sensoriel, la complexité du traitement auditif est évidente puisque cet organe reçoit également des entrées efférentes depuis le tronc cérébral. Le circuit auditif ascendant transmet l'information le long du tronc cérébral (du noyau cochléaire au complexe de l'olive supérieure et jusqu'au colliculus inférieur) et du thalamus (corps genouillé médian) avant de transmettre l'information au cortex auditif primaire. L'information dans le cortex diverge ensuite entre les circuits auditifs ventral et dorsal, d'une façon similaire à la vision, en se basant sur les traitements du "quoi" et du "où". Il a été montré que les aires sous-corticales du système auditif - corps géniculé médian et colliculus inférieur – font preuve de plasticité, les impliquant dans la mémoire. Toutefois sont-elles impliquées dans la mémoire auditive pour les sons dénués de sens ?

Depuis la découverte par Guttman et Julesz que les adultes sont capables de détecter et de conserver une information auditive sans signification dans leur mémoire de travail, de nombreuses expériences ont étudié les propriétés de cette capacité: (i) la capacité à conserver ces sons a été explorée chez différents mammifères; (ii) les réponses comportementales aux sons cycliques ont été étudiées en utilisant un tâche cyclique de « tapping »; (iii) une autre série d'expériences a utilisé ce paradigme pour comprendre les limites et les capacités de la perception auditive et de la mémoire de travail et enfin ; (iv) en utilisant un paradigme expérimental élégant, Agus, Thorpe et Pressnitzer ont été les premiers à étudier la mémoire sensorielle à long-terme en utilisant des sons cycliques (Agus et al., 2010).

En utilisant un paradigme d'apprentissage implicite où plusieurs présentations de quelques sons cycliques (sons cycliques cibles) sont présentées, Agus et al. ont montré que les participants peuvent stocker cette information de façon robuste pendant plusieurs semaines. Les corrélats neuraux de cette mémoire ont ensuite été étudiés en EEG (Andrillon et al., 2015; Luo et al., 2013) et IRMf (Kumar et al., 2014). L'ensemble des résultats de ces expériences de comportement et de neuroimagerie a soulevé de nombreuses questions quant aux moyens et à la localisation de la perception et du stockage de ces propriétés acoustiques dans le cerveau. Durant les trois années de ma thèse, je me suis attachée à adresser certaines de ces questions dans la série d'expériences suivantes :

Expérience I: Robustesse de la mémoire pour les sons gaussiens implicitement appris.

Dans cette expérience, nous avons exploré la capacité des participants à se rappeler des versions modifiées d'un bruit gaussien appris en utilisant un dispositif expérimental similaire à celui utilisé par Agus et al. Nous souhaitions en effet tester la robustesse de la mémoire de reconnaissance des participants pour ces sons lorsque ceux-ci leurs sont présentés dans des versions dégradées en comparaison de la version apprise. Pour cela, les versions « looped » et « scrambled » (brouillées par segment de 10 ou 20 ms) d'un son cyclique appris 4 semaines avant le test ont été présentées aux participants.

Résultats: Prolongeant les résultats précédemment obtenus par Agus et al., nous avons démontré que les participants peuvent stocker des informations implicitement apprises jusqu'à 4 semaines. Ils détectent mieux les versions intactes, looped et scrambled d'un son cyclique déjà appris par rapport à de nouveaux sons cycliques suggérant que les neurones peuvent encoder de très petits fragments d'information dans la modalité auditive. De façon intéressante, la performance durant la session d'apprentissage a dicté la performance en mémoire de reconnaissance pour les versions intactes et dégradées des sons cibles.

Modèle STDP: Les sons utilisés dans l'expérience précédente ont été implémentés dans un modèle STDP (Masquelier et al., 2016). Ainsi, comme avec les participants, les versions intactes, looped et scrambled d'un son cyclique, ainsi qu'un nouveau son cyclique ont été présentés au neurone LIF et les patterns des potentiels d'action en sortie ont été comparés. Les sorties obtenues avec notre modèle sont en parfait accord avec les résultats comportementaux obtenus.

Expérience II: Corrélats spatiaux de la mémoire pour les sons gaussiens implicitement appris.

Dans cette expérience, nous avons étudié les corrélats neuraux de la mémoire pour les sons dénués de sens en utilisant un paradigme expérimental en IRMf. Plus spécifiquement, nous avons testé l'hypothèse que les aires sous-corticales impliquées dans la perception de cette information sont également impliquées dans son stockage. Nous avons également fait l'hypothèse que l'hippocampe et les aires du cortex auditif sont impliqués dans le stockage des propriétés acoustiques. La session d'apprentissage était similaire à celle décrite par Agus et al. et précédemment utilisée dans notre expérience I. La tâche des participants consistait à discriminer les sons cycliques des non-cycliques lors de la session de test réalisée en IRMf. Les « meilleurs » et « pires » sons cycliques identifiés lors de la session d'apprentissage ainsi que de nouveaux sons ont été présentés aux participants.

Résultats: Les participants discriminent préférentiellement les sons cycliques précédemment appris (meilleurs discrimination par rapport aux nouveaux sons). Les contrastes fonctionnels obtenus en IRMf impliquent des zones inférieures de la voie auditive (colliculus inférieur et le corps genouillé médian) de même que l'hippocampe dans l'établissement de cette mémoire.

Expérience III: Etude des mécanismes et limites de résolution de la mémoire pour les sons gaussiens.

La capacité des participants à mémoriser des sons de 10 ms brouillés constituait un résultat surprenant ; c'est pourquoi nous avons voulu explorer les mécanismes neuraux sous-jacents. Pour cela nous avons mis en place un paradigme expérimental nouveau et inhabituel, dans lequel des sons cycliques (de différentes fréquences) et un bruit plein étaient présentés dans chaque oreille des participants. Il était demandé aux participants d'identifier la source du son cyclique. Implicitement, certains sons cycliques étaient présentés à plusieurs reprises.

La mémoire pour ces segments de bruit implicitement encodés a été testée après 4 semaines dans une série d'expériences de mémoire implicite et explicite. Résultats: Bien que les participants n'aient pas de mémoire consciente des sons appris, ils sont plus performants dans la tâche de localisation de sons cycliques de 10 ms appris que les sons cycliques nouveaux, dès 8 répétitions (soit 80 ms). Nous avons également mis en évidence une relation linéaire entre le nombre de présentations et l'efficacité du rappel.

Conclusions: L'ensemble de nos résultats met en évidence la robustesse de la mémoire pour les sons dénués de sens sans aucun accès conscient, et ses mécanismes neuraux sous-jacents. La performance en mémoire de reconnaissance est dépendante de la sensibilité des participants aux différences de propriétés acoustiques des bruits (a') et non pas de la performance sur la tâche d'apprentissage (sons cycliques les « mieux » appris ou les « moins bien » appris). Une relation linéaire entre nombre de présentation et efficacité de la reconnaissance a été observée. Aucun accès conscient à cette information n'a été observé dans la tâche de localisation du son impliquant le circuit auditif dorsal (traitement inconscient de l'information "où") de même que la tache de discrimination cyclique/non-cyclique impliquant la voie auditive ventrale (traitement conscient de l'information « quoi »), peut-être du fait d'un stockage sous-cortical des propriétés de l'information. Cette mémoire de reconnaissance implicite apparait également être robuste et résistante à l'altération. L'ensemble de ces résultats conforte les prédictions obtenues avec le modèle STDP d'apprentissage de motifs répétés de bruit; suggérant que la STDP constitue le mécanisme cellulaire candidat dans le stockage de l'information sensorielle.

L'ensemble de ces expériences met en évidence deux résultats surprenants : 1) la capacité d'apprentissage et de stockage de motifs acoustiques de moins de 10 ms et 2) la localisation de cette mémoire qui semble être sous-corticale, et qui survient dans des régions similaires à celles impliquées dans la perception des sons.

Humans are able to detect acoustic features in Gaussian noise. Researchers recently used repeating noise segments [*cyclic noises (CNs)*, presenting a segment of noise several times back to back] to investigate long-term sensory memory (Agus et al., 2010). They asked participants to discriminate CNs from plain noise, while implicitly presenting them with a few *target* CNs several times. The results demonstrated long-term memory for such sounds, which have raised several further questions. First, the *robustness of memory* for implicitly learned Gaussian sounds was tested using a similar paradigm. Participants' recognition memory was tested by presenting them with looped and scrambled (10 or 20-ms bin size) versions of target CNs 4 weeks post-learning. Our results suggest that neurons might code for very small bits of acoustic information (10 ms). Next, the *spatial correlates of memory*, specifically, the role of subcortical areas in storing auditory patterns was investigated. Using the same paradigm, participants performed the testing session during fMRI scanning. Implicit memory for target CNs was demonstrated and functional contrasts implicate the Medial Geniculate body and hippocampus. Lastly, we explored the *mechanisms and resolution limits of this memory*. Participants were presented with CNs in one ear and plain noise in the other ear, and had to localize the CN. Implicit and explicit memory for target CNs was tested 4 weeks later. Although participants lacked conscious memory, they were better at localizing target 10-ms CNs than novel CNs, even with 8 repeats (80 ms). Altogether we demonstrate: 1) the ability to learn and *store short acoustic patterns (10 ms)*; 2) this *memory is sub-cortical*, in regions implicated in *perception* of sounds; and 3) these results are compatible with an *STDP model of learning*.

Key words: Long-term memory, resolution of auditory representations, sensory learning, fMRI, implicit memory, explicit memory, Gaussian sounds.

L'homme peut discriminer les caractéristiques acoustiques de bruits Gaussiens. Les mécanismes de la mémoire sensorielle à long terme ont récemment été étudiés en utilisant des segments de bruit répétés en continu, ou bruits cycliques (CNs) (Agus et al., 2010). Les sujets devaient discriminer des CNs d'autres bruits aléatoires, certains CNs *cibles* étant présentés plusieurs fois à l'insu des sujets. Une mémorisation à long terme de ces CNs cibles a été démontrée, soulevant des questions quant aux mécanismes mnésiques sous-jacents. Ici, nous avons étudié *la robustesse de cette mémoire*, en testant la reconnaissance implicite à long terme (1 mois) de CNs cibles transformés : son enroulé sur lui-même (CNs « looped »), ou brouillé (CNs « scrambled », 10 ou 20 ms). Nous montrons que de très courts segments de bruit peuvent être stockés en mémoire à long terme (10 ms). *Le rôle des structures* (sous-corticales) *dans cette reconnaissance* à long terme a ensuite été étudié par IRMf. Nous observons une trace mnésique des CNs cibles impliquant les premiers relais de la voie auditive, en particulier le corps genouillé médian, ainsi que l'hippocampe. Enfin, nous avons exploré *les limites de cette mémoire* en présentant des CNs cibles de différentes durées dans une oreille, et des bruits purement aléatoires dans l'autre oreille ; les sujets devant localiser le CN. Un mois après, les sujets ont une reconnaissance implicite de CNs cibles aussi brefs que 10 ms, avec seulement 8 répétitions (80ms). Nous démontrons ainsi : 1) la capacité d'apprendre et de conserver en mémoire des segments de bruit aussi *courts que 10 ms*, 2) une *trace mnésique sous-corticale*, dans les régions impliqués dans *la perception* des sons, 3) ces résultats sont en accord avec les performances de reconnaissance prédites par *un modèle d'apprentissage STDP*.

Mots-clés : Mémoire à long terme, résolution des représentations auditives, apprentissage sensoriel, IRMf, mémoire explicite, mémoire implicite, sons gaussiens.