Conscious and unconscious processing of temporal regularities : a joint modeling and experimental approach

Catherine Wacongne

Université Pierre et Marie Curie

Ecole Doctorale Cerveau, Cognition, Comportement

THESE

présentée pour obtenir le grade de

DOCTEUR EN SCIENCES

Spécialité : NEUROSCIENCES COGNITIVES

TRAITEMENTS CONSCIENT ET NON-CONSCIENT DES REGULARITES TEMPORELLES

MODELISATION ET NEUROIMAGERIE

Par Catherine WACONGNE

Dirigée par Stanislas DEHAENE

Et Co-encadrée par Jean-Pierre CHANGEUX

Jury:

Floris DE LANGE	Rapporteur
Karl FRISTON	Rapporteur
Gustavo DECO	Examinateur
Lionel NACCACHE	Examinateur
Jean-Pierre CHANGEUX	Examinateur
Stanislas DEHAENE	Examinateur

Université Pierre et Marie Curie

Ecole Doctorale Cerveau, Cognition, Comportement

Dissertation submitted for the degree of

Doctor of Philosophy

COGNITIVE NEUROSCIENCE

CONSCIOUS AND UNCONSCIOUS PROCESSING OF TEMPORAL REGULARITIES

A JOINT MODELING AND EXPERIMENTAL APPROACH

By Catherine WACONGNE

Supervised by Stanislas DEHAENE

And Co-supervised par Jean-Pierre CHANGEUX

Jury:

Floris DE LANGE	Rapporteur
Karl FRISTON	Rapporteur
Gustavo DECO	Examiner
Lionel NACCACHE	Examiner
Jean-Pierre CHANGEUX	Examiner
Stanislas DEHAENE	Examiner

TABLE OF CONTENTS

CHAPTER1: INTRODUCTION 12

1.1 USING EVENT RELATED POTENTIALS TO UNEXPECTED EVENTS TO ASSESS CONSCIOUSNESS IN NON-COMMUNICATIVE PATIENTS 12

1.1.1 ASSESSING CONSCIOUSNESS IN NON-COMMUNICATIVE PATIENTS 12

1.1.2 CONSCIOUSNESS: THEORIES OF CONSCIOUSNESS AND NEURONAL CORRELATES OF CONSCIOUS ACCESS 14

1.1.2.1 Theories of consciousness 15

1.1.2.2 Neuronal correlates of conscious access in MEEG. 18

1.1.3 THE "LOCAL-GLOBAL" PARADIGM 22

1.1.4 MODELING CHALLENGES 24

1.2 WHY DO WE PROCESS TEMPORAL REGULARITIES? 25

1.2.1 PREDICTIVE CODING 25

1.2.2 BEHAVIORAL CONSEQUENCES OF TEMPORAL REGULARITY LEARNING 33

1.2.2.1 Reward anticipation and decision making 33

1.2.2.2 Motor control optimization 37

1.2.2.3 Orienting of attention towards new events 38

1.2.3 FREE-ENERGY PRINCIPLE 40

1.3 WHAT KIND OF REGULARITIES CAN WE LEARN? 41

1.3.1 CHOMSKY HIERARCHY: A FRAMEWORK FOR THE CLASSIFICATION OF SEQUENTIAL REGULARITIES 42

1.3.1.1 Formal language theory (FLT) 42

1.3.1.2 Human natural languages and the Chomsky hierarchy 47

1.3.1.3 Using and learning Chomsky hierarchy 49

1.3.1.4 Multiple levels of complexity: an evidence for a modular processing of temporal regularity? 50

1.3.2 EXPERIMENTAL APPROACHES TO CONSCIOUS AND UNCONSCIOUS REGULARITY PROCESSING 51

1.3.2.1 Temporal regularity learning and awareness of the rule 51

1.3.2.2 Temporal regularity learning and central resources 55

1.3.3 SIMULTANEOUS BUT OPPOSITE EXPECTATIONS FOR CONSCIOUS AND UNCONSCIOUS PROCESSING 57

1.4 NEURONAL PROPERTIES FOR THE IMPLEMENTATION OF A TEMPORAL PREDICTIVE PROCESS. 58

1.4.1 TEMPORAL DYNAMICS OF LEARNING 58

1.4.2 NEURONAL MECHANISMS OF SEQUENCE LEARNING 60

1.4.2.1 Short term plasticity 60

1.4.2.2 Long term plasticity 61

1.4.3 NEURONAL CODES IN AUDITORY CORTEX 64

1.4.4 PREDICTING TIMING AND IDENTITY IN NEURONAL NETWORKS 65

2 CHAPTER2: INTRODUCTION TO THE PERSONAL CONTRIBUTIONS 72

3 CHAPTER3: ARTICLE 1: A NEURONAL MODEL OF PREDICTIVE CODING ACCOUNTING FOR THE MISMATCH NEGATIVITY 76

3.1 INTRODUCTION TO THE ARTICLE 76

3.1.1 GOAL OF THE ARTICLE 76

3.1.2 CHOICE OF THE METHODS 76

3.1.3 REFERENCE 77

3.2 ARTICLE 78

3.2.1 ABSTRACT 78

3.2.2 INTRODUCTION 78

3.2.3 MATERIAL AND METHODS 80

3.2.3.1 Network architecture 81

3.2.3.2 Detailed implementation 82

3.2.3.3 Spiking neuron model 82

3.2.3.4 Synaptic plasticity 84

3.2.3.5 Simulations: 85

3.2.3.6 MEG experiment 85

3.2.4 RESULTS 88

3.2.4.1 Oddball paradigm and MMN 88

3.2.4.2 Behavior of the memory neurons 90

3.2.4.3 92

3.2.4.4 Layer distribution of current sources. 92

3.2.4.5 Effect of Deviant probability. 92

3.2.4.6 Internal model of the temporal statistics in the input 93

3.2.4.7 MMN to repetition in an alternate signal 95

3.2.4.8 Blindness to global regularities 96

3.2.4.9 MMN to omission 97

3.2.4.10 MMN to changes in duration 99

3.2.4.11 Prediction vs. habituation : an experimental test of the model 100

3.2.5 DISCUSSION 104

3.2.5.1 Predictions versus synaptic habituation 104

3.2.5.2 Extensions and limits of the model 105

3.2.6 CONCLUSION 107

3.2.7 ACKNOWLEDGEMENT 107

4 CHAPTER 4: ARTICLE 2: EVIDENCE FOR A HIERARCHY OF PREDICTIONS AND PREDICTION ERRORS IN HUMAN CORTEX 110

4.1 INTRODUCTION TO THE ARTICLE 110

4.1.1 GOAL OF THE STUDY 110

4.1.2 REFERENCE 110

4.2 ARTICLE 111

4.2.1 ABSTRACT 111

4.2.2 INTRODUCTION 112

4.2.3 METHODS 115

4.2.3.1 Subjects 115

4.2.3.2 Auditory Stimuli 115

4.2.3.3 Simultaneous EEE-MEG recordings 115

4.2.3.4 Data analysis 116

4.2.3.5 Source reconstruction 116

4.2.4 RESULTS 117

4.2.5 DISCUSSION 121

4.2.6 ACKNOWLEGMENTS 125

5 CHAPTER 5: MODELING ACCESS TO WORKING MEMORY AS A SELF-EVALUATION AND DECISION PROCESS 128

5.1 INTRODUCTION TO THE ARTICLE 128

5.2 ARTICLE 129

5.2.1 ABSTRACT 129

5.2.2 INTRODUCTION 129

5.2.3 METHODS 130

5.2.3.1 General notations 130

5.2.3.2 Model description 131

5.2.3.3 Initial conditions 132

5.2.3.4 Performance estimate 132

5.2.4 RESULTS 133

5.2.4.1 Simplification hypothesis 133

5.2.4.2 136

5.2.4.3 Ability to discover relevant dependencies 137

5.2.4.4 Ability to develop the right strategy 138

5.2.4.5 Dependency between speed of rule discovery and divergence between the probability distributions 141

5.2.4.6 Comparison with other models *142*
5.2.4.7 Learning of more complex sequential regularities *143*
5.2.4.8 Learning of the Local-global paradigm *145*
5.2.5 DISCUSSION 145
6 CHAPTER 6: GENERAL DISCUSSION 152
6.1 SUMMARY OF THE MAIN RESULTS 152
6.2 DISCUSSION 154
6.2.1 GENERAL PRINCIPLES FOR TEMPORAL REGULARITY LEARNING 155
6.2.1.1 Roles of predictions and prediction errors *155*
6.2.1.2 Learning implicit models of the world *155*
6.2.2 SPECIFICITY OF CONSCIOUS PROCESSING 156
6.2.2.1 Exact timing and event based timing *157*
6.2.2.2 Abstract rules *157*
6.3 CONCLUSION 162
BIBLIOGRAPHY 164

TABLE OF FIGURES

Figure 1.1-1: Consciousness and the global neuronal workspace theory 16
Figure 1.1-2: Schematic representation of 3 models of MMN 21
Figure 1.1-3: The "local global" paradigm. 23
Figure 1.2-1: Internal models. 27
Figure 1.2-2: Canonical cortical microcircuit and predictive coding 30
Figure 1.2-3: Hierarchical predictive coding and global neuronal workspace architectures 32
Figure 1.2-4: Response of Dopamine Neurons to Reward and predicted reward 35
Figure 1.3-1: The hierarchy of grammars in Formal Language Theory 46
Figure 1.3-2: Trace and delay conditioning. 55
Figure 1.4-1 Effect of recent history on temporal regularity processing 60
Figure 1.4-2: Spike timing dependent plasticity 63
Figure 1.4-3: Schematic representation of the main models of working memory 68
Figure 3.2-1: Scheme of the predictive coding model for two sounds 80
Figure 3.2-2 Simulating the MMN in an oddball paradigm : mean synaptic currents and firing rates. 89
Figure 3.2-3: Simulated pattern of neural firing and membrane voltage during a single trial of the oddball paradigm. 91
Figure 3.2-4: Correspondence between the transition statistics of the inputs (left) and the synaptic weights learned by the model (right). 93
Figure 3.2-5: Simulating the MMN in response to an unexpected repetition amongst alternating stimuli. 95
Figure 3.2-6: Simulating the lack of sensitivity of the MMN to global regularities that cannot be captured by local transition statistics. 97
Figure 3.2-7: Simulating the MMN to the omission of an expected sound. 98
Figure 3.2-8: Simulating the MMN to a duration deviant. 99
Figure 3.2-9: Experimental test of the model using magneto-encephalography. 102
Figure 4.2-1 : EXPERIMENTAL DESIGN. 113
Figure 4.2-2: Sensor-level topography and time course of the brain responses to distinct forms of novelty. 118
Figure 4.2-3: Source modeling of the effects. 120
Figure 5.2-1: The simplified computational problem studied. 133
Figure 5.2-2: Comparison between classical reinforcement learning 135
Figure 5.2-3: Functioning of the model at each time step. 136
Figure 5.2-4The model discovers hidden regularities even at a long temporal distance. 138
Figure 5.2-5: Dynamics of discovery of a successful policy. 139

Figure 5.2-6: A lawful relationship relates the time of successful policy discovery to the informativeness of the predictive stimulus. *141*

Figure 5.2-7: The present model achieves a good compromise between final performance and speed of learning, especially for long distance predictive relations. *142*

Figure 5.2-8: Variability in the policy discovered on 20 different runs of the same model. *144*

Figure 6.2-1 : explicit representation of an abstract rule in lateral prefrontal cortex (lPFC). *158*

CHAPTER1

INTRODUCTION

Temporal regularities are present everywhere: in the position of objects across time, in the vocalization of animals, in the ticking of a clock… The rules of causality impose that similar causes have reproducible effects. As a result, events tend to occur in a reproducible order. Being able to predict what is going to happen next is obviously an asset for survival. In this thesis, I will try to better understand how the brain takes advantage of temporally regular structures using a modeling approach, combined with neuroimaging experiments to test specific theory-driven hypothesis. In particular, I will focus on understanding the role of conscious processing in temporal regularity learning.

In this introductive chapter, I will first introduce in more details what conscious processing is and why I chose to focus on its role in temporal regularity learning. Following David Marr's hierarchy (Marr, 1982) , I will then review the main relevant frameworks that were used to understand the computational goals of regularity learning. Next, I will try to establish what characterizes at the algorithmic level the type of processes involved in conscious and unconscious temporal regularity processing. Finally, I will present some of the neuronal mechanisms that are relevant to understand how temporal regularity learning might be implemented in the brain.

1.1 Using event related potentials to unexpected events to assess consciousness in non-communicative patients

1.1.1 Assessing consciousness in non-communicative patients

For most of us, arousal and awareness are two closely correlated concepts. When we go to sleep or undergo global anesthesia, we lose both; when we awake, both are recovered at the same time. This also holds for most patients unlucky enough to go into a coma; however, there are a few patients that show a strange dissociation. One of the first case was described by Rosenblath (Rosenblath, 1899). A young tightrope-walker had fallen from his wire into a coma. Two weeks after the accident, the patient became "strangely awake": he presented clear signs of

arousal but was still unable to interact meaningfully with its environment. The patient died a few weeks after from its injuries without recovering further. Nowadays, we use the term "vegetative state" to describe these patients, because they are considered "to live a merely physical life devoid of intellectual activity […] social intercourse [...] sensation and thought" (Jennett & Plum, 1972, p 736). Clinically, they are characterized as patients that present signs of arousal without awareness. People in a vegetative state may open their eyes, wake up and fall asleep at regular intervals and may have basic reflexes such as blinking when they are startled by a loud noise or withdrawing their hand when it's squeezed hard. They are also able to regulate their heartbeat and breath without assistance. However, patients in a vegetative state do not show any meaningful responses to external stimulations such as following an object with their eyes or responding to surrounding voices. They will also not show any sign of experiencing emotions.

The diagnosis of vegetative state – arousal without awareness – is a challenging problem. Although arousal has a clear behavioral definition, it is less straightforward to provide convincing proof of the absence of awareness. The current diagnosis criteria involve the inability to follow simple commands or to communicate verbally. Yet, the inability to communicate does not necessarily imply the absence of awareness. Patients presenting the so called "locked-in" syndrome are the typical example of this dissociation (Plum & Posner, 1966). They are most often completely paralyzed, except for vertical eye movements and blinking. In extreme cases, called total locked-in syndrome, even eye movements are impossible and no communicative behavior can therefore be observed (Bauer, Gerstenbrand, & Rumpl, 1979). These patients can nonetheless be perfectly aware of their environment.

As a result, the possibility that some vegetative state patients could be misdiagnosed must be considered. How can we objectively assess something as subjective as consciousness? How can we know whether patients experience their surrounding environment when they are utterly unresponsive?

It is necessary to overcome the limitations of behavioral measures to test the hypothesis that consciousness might be intact while its behavioral expression is impossible. A groundbreaking study (A. M. Owen et al., 2006) proposed to solve this issue by using a more direct communication channel: functional Magnetic Resonance Imaging (fMRI) activations. They first asked healthy volunteers to perform two tasks of mental imagery. The first task consisted in imagining playing tennis. It recruited the Supplementary Motor Area (SMA) and subjects showed an increase in the BOLD signal in this area. The second task consisted in imagining navigating in a room of their house, which activated the para-hippocampal gyrus, the posterior parietal cortex

and the lateral premotor cortex. Then the author applied the same protocol to patients diagnosed as vegetative. The decisive result was that one of these patients showed activations that were similar to the ones of healthy subjects, thus showing his ability to understand complex instruction and voluntarily modulate his brain activity by accomplishing a mental task. The results of this study were reproduced in an other group of patients, where five out of fifty-four vegetative state patients showed voluntary modulation of their brain activations according to the instructions (Monti et al., 2010). One of them was even able to answer yes-or-no questions using the same mental imagery tasks.

This approach was very promising but it relies on fMRI, which is a powerful but extremely expensive imaging technique that requires transporting patients that are quite fragile and may present contraindications. Therefore, it is hardly realistic to imagine that this protocol could be used as a routine diagnostic tool in clinics, let alone used multiple times in the same patients at different stages of their potential recovery. It is therefore necessary to devise protocols using cheaper imaging techniques that can be used at the patient bedside. Several attempts to adapt the protocol to EEG (Cruse et al., 2011; A. M. Goldfine, Victor, Conte, Bardin, & Schiff, 2011) were made but subjected to controversies about the validity of the statistical analysis involved (Cruse et al., 2013; A. Goldfine et al., 2013).

1.1.2 Consciousness: theories of consciousness and neuronal correlates of conscious access

At this point, it is necessary to clarify the definition of consciousness. Indeed, consciousness can be an ambiguous concept. In its intransitive form ("The patient is still conscious"), it refers to the general *state* of consciousness also called wakefulness or vigilance and can vary in a graded manner from coma to full vigilance. In its intransitive form ("I was not conscious of the car coming to my right") (Stanislas Dehaene & Changeux, 2011) it refers to the conscious *access* or *processing* of specific piece of information. Conscious access is generally determined by the ability to report verbally the information being accessed.

Although it is necessary to *be* conscious to be conscious *of* something, the reverse is not necessarily true. Theories of consciousness have mainly focused on the second aspect: what determines whether we are conscious or not of a stimulus? What are the neuronal correlates of conscious access?

1.1.2.1 Theories of consciousness

The word "consciousness" has long been banned in cognitive science, following the behaviorist tradition that tried to banish subjectivity from the study of cognition. Therefore, the first models of selective processes in perception were not referring explicitly to consciousness, but rather to "working memory" or "selective attention". A simple but influential model of selective central processing was proposed by Broadbent (Broadbent, 1957) and dissociated two steps of processing: a first step of processing in parallel sensory buffers, followed by a unique "perceptual" system with limited capacity which would explain why we cannot perceive multiple things at the same time. This distinction between early parallel processors and a selective central system dependent on attention where a percept can be maintained in a more permanent manner was used in other early models (Baddeley & Hitch, 1974; Norman & Shallice, 1986). It is now the basis of the more contemporary models of conscious access.

An important issue of these models was the presence of a central, somewhat obscurely defined controller which seemed to correspond to an external observer that would supervise the rest of the system. Consciousness itself was described using a theater where that narrow stage would only allow for one actor to diffuse its message to some spectator (Taine, 1882). To overcome the recursive problem of explaining how this "homunculus" would be itself conscious (Dennett, 1992), current theories of consciousness now prefer distributed architectures that seem to escape the problem of the internal observer by presenting a large capacity to represent and route information. One of these models is the "global workspace" theory (B. Baars, 1989). It postulates a large number of parallel input processors that compete for a capacity-limited central space, which itself can broadcast its content to the input processors. According to this theory, the information of the processor can be processed either consciously, if it gained access to the central space, or unconsciously if it did not.

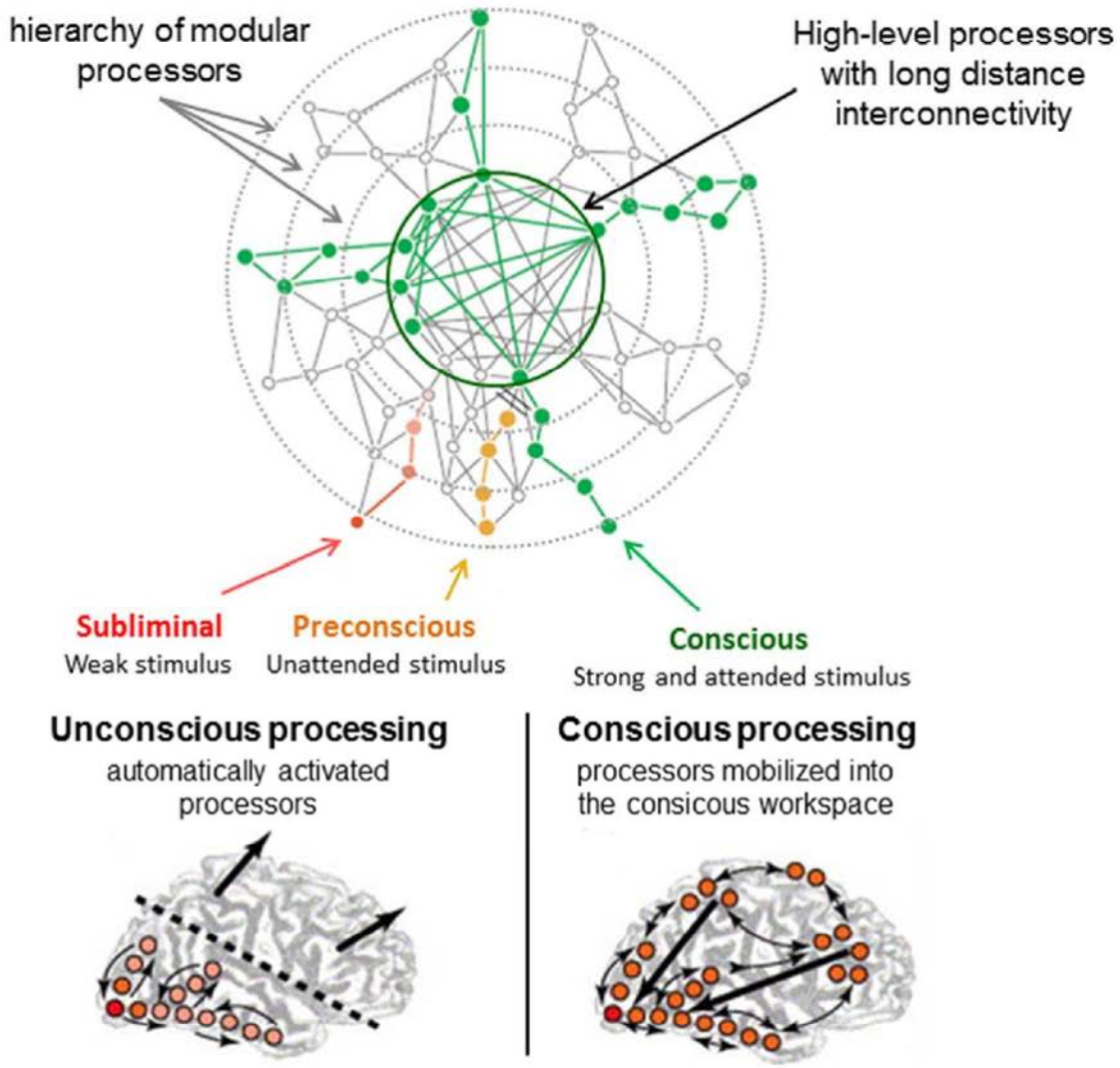


FIGURE 1.1-1: CONSCIOUSNESS AND THE GLOBAL NEURONAL WORKSPACE THEORY

(Top) Schematic representation of the neuronal architecture underlying the globa neuronal workspace model. Stimuli that are too week or unattended are processed by a hierarchy of automatic cortical processors. If a stimulus is sufficiently strong and attended, it reaches high level processors that are highly interconnected via long distance connections. The information that reaches this central space can be broadcasted to other processors in a top down manner. (Bottom) Conscious and unconscious processing: unconscious stimuli do not reach the frontopariaetal network. The stimulus related activity remains restricted to sensory areas. Conscious stimuli reach the frontoparietal network with a phenomenon called ignition, where the information is actively maintained for a more extended period of time. Adapted from (Stanislas Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006)

These first "psychological" models of consciousness paved the way for more neurobiological approaches. My goal here is not to provide an exhaustive review of neurobiological theories of consciousness, but rather to point towards neuronal architectures and

properties that have been proposed to play an important role in conscious access (see Seth, 2007 for a more complete review).

- First, although some theories argue that consciousness might not be a unitary process and can be localized in particular regions (Zeki, 2003), most theories make the hypothesis that conscious access is associated with a winner-take-all competition for a *central resource* with *limited capacity* that relies on a distributed architecture. This distributed architecture involves in many models the prefrontal cortex (PFC) (Stanislas Dehaene, Kerszberg, & Changeux, 1998b; Lau, 2008; Norman & Shallice, 1986; Posner & Rothbart, 1998) and the executive and selective functions that are associated with it.
- Second, consciousness is mainly thought as an integrative process, involving *recurrent* or *re-entrant* connections or *feedback* processing. In Edelman's theory (Edelman, 1989; Sporns, Tononi, & Edelman, 2000) the conscious state relies a re-entrant connectivity involving the cortico-thalamic loop, that produces representations that are both highly differentiated (there can be a large number of representations) and highly integrated (the information is shared across a large set of regions). Lamme (Lamme, 2010) insists on the importance of recurrent processing in the cortex for conscious experience. He distinguished two stages in perception: the first is a fast and mainly feedforward sweep of neuronal activity, followed by recurrent processing (Lamme & Roelfsema, 2000) that would correspond to conscious perception. For example, visual information would first be processed by the two pathways of visual cortex, leading to the identification of the different features of the input (color, motion, shape…). This information can then be maintained, shared and integrated via horizontal and feedback connections. Finally, Dehaene and Changeux (Stanislas Dehaene, Kerszberg, & Changeux, 1998a) proposed a "global neuronal workspace" theory (Figure 1.1-1), in which the automatic sensory processor operate in parallel and compete for access to the global workspace. This workspace is characterized by a network of pyramidal neurons with long range connectivity that allows *maintenance* of information (see also B. J. Baars & Franklin, 2003) through a recurrent connectivity, and the diffusion of the centrally maintained information to lower level areas.

1.1.2.2 Neuronal correlates of conscious access in MEEG.

The theories presented above are based on experimental data that investigated the neuronal correlates of conscious access. As argued before, we will focus here on a specific imaging technique: the electroencephalography (EEG). The EEG was used for the first time in humans by Hans Berger in 1924 (Berger, 1929) and simply relies on electrodes disposed on the head to record a summation of the electrical activity at the surface of the scalp. The main sources of EEG potentials are thought to be the synaptic currents generated when large populations of neurons receive similar synaptic stimulations. Because EEG is a continuous measure that follows directly the electrical activity of neurons, it has an excellent temporal resolution and as a consequence, is a privileged technique over coarser temporally resolved methods such as fMRI, to investigate temporal regularity processing. For forty years, EEG was used a continuous measure providing criteria to identify states of sleep or arousal, or study epileptic activity. It was only in 1964 that the first evoked potentials (ERP) were identified by averaging many epochs of signal evoked by multiple presentations of a stimulus (Walter, Cooper, Aldridge, McCallum, & Winter, 1964).

Consistently with the exceptional power of ERPs to resolve temporal steps of processing, some of the first studies exploiting this new technique of analysis tried to develop paradigms involving minimally different conditions that would highlight different stages of processing. The P300 was one of the first potentials identified this way (Chapman & Bragdon, 1964), and was related early on to the study of conscious access (Desmedt, Debrecker, & Manil, 1965) but also to temporal regularity processing (S. Sutton, Braren, Zubin, & John, 1965). Desmedt used an auditory cue to prime a faint tactile stimulus. He observed that when the stimulus was cued, it was more often detected. Looking at the potentials evoked by cued and uncued stimuli, he observed that the early ERP (<300ms after the onset of the tactile stimulus) were identical in both conditions, whereas the late ERP differed. Unfortunately, the authors asked the subjects to count the stimuli they detected, so the hypothesis that the late potential was due to the counting activity and not the conscious detection *per se* could not be ruled out. A few years later, Hillyard (Hillyard et al., 1971) used auditory stimuli at perceptual threshold, and asked subjects to report whether they perceived the stimulus using delayed motor response. He showed that the early potentials were unaffected by the detection performance, while the P300 was only present when the subjects detected the stimulus.

The P300 was later dissociated in two components called P3a and P3b (N. Squires, Squires, & Hillyard, 1975). P3a and P3b can be decorrelated using a three stimulus oddball

paradigm: one frequent sound (the standard) is repeated at regular intervals and is sometimes replaced by one of two possible deviants. One of these deviants is very similar to the standard tone. It is used as the target of a difficult detection task. The other deviant is very different from the standard, and is used as a distractor. In this case, the distractor stimulus elicits only a P3a, while the target, task relevant stimulus, elicits only a P3b. The P3a is the first subcomponent, observed transiently over the frontal electrodes and elicited by unexpected events (K. C. Squires, Wickens, Squires, & Donchin, 1976; S. Sutton et al., 1965). The P3b is a more sustained component over the centro-posterior electrodes and is only elicited by visible, task relevant stimuli (Stanislas Dehaene et al., 2006; Stanislas Dehaene & Changeux, 2011; Polich & Criado, 2006; Polich, 2007; N. Squires et al., 1975). A variety of experimental manipulations have shown that the P3b is likely to be the best a good index of conscious detection in EEG and magnetoencephalography (MEG) (Del Cul, Baillet, & Dehaene, 2007; Gutschalk, Micheyl, & Oxenham, 2008; Sekar, Findley, Poeppel, & Llinas, 2013; Sergent, Baillet, & Dehaene, 2005; van Aalderen-Smeets, Oostenveld, & Schwarzbach, 2006).

In contrast, the early potentials are less sensitive to attentional or consciousness manipulations. The mismatch negativity (MMN) is characteristic of this automatic processing of information. Although it was already observed along with the P300 in previous studies (K. C. Squires et al., 1976), the MMN was first distinguished from the other attention-dependent potentials by Näätänen (Näätänen, Gaillard, & Mäntysalo, 1978). The MMN is a negative component of the ERP elicited by any perceptible change in some repetitive aspect of auditory stimulation. It is typically elicited in an oddball paradigm, where a frequent tone (the standard) is repeated at regular intervals. Rarely, a different tone is presented instead (the deviant). The roles of the two tones can be swapped in a different block. The difference between the waveforms evoked by a given sound presented as the frequent and as the rare stimulus is an early negative ERP component that is maximal over frontocentral areas of the scalp and peaks between 100-200ms. The existence of an MMN is independent of attention (Alain, Woods, & Ogawa, 1994; Alho, Sams, Paavilainen, Reinikainen, & Näätänen, 1989; Alho, Woods, & Algazi, 1994; Muller-Gass, Stelmack, & Campbell, 2005; Näätänen et al., 1978), and its amplitude is modulated very little by attention although it can be slightly attenuated (Muller-Gass et al., 2005). A significant MMN can also be found under anesthesia (Heinke et al., 2004; Koelsch, Heinke, Sammler, & Olthoff, 2006; Simpson et al., 2002). Therefore it is thought to be the result of a low-level automatic processing of auditory inputs. The dominant interpretation of MMN (Figure 1.1-2) is that it is the result of a comparison process between a sensory memory (echoic memory) that encodes the repetitive aspects of the stimulus, and the incoming input (Näätänen, 2003). MMN

originates mainly form primary auditory cortex. For many years, MMN was considered to be elicited by a separate source of another ERP occurring in the same time window after the onset of the stimulus: the N1 component. The co-existence of two separable components at the same time was interpreted as evidence for two separate processing pathways: the N1 would reflect general processing of the information, and the MMN would reflect a separate mechanism dedicated to the detection of inconsistency between the context established by previous stimuli and the incoming input. However, more recent analysis suggest that MMN could be a modulation of the N1 response instead, making unnecessary the hypothesis of multiple routes (Ahveninen et al., 2004). An alternative "fresh afferent" or "adaptation" model interprets MMN as reflecting short term synaptic depression of synapses that are repeatedly stimulated. When a different sound is presented, it would elicit a bigger response by stimulating new synapses. This model is largely supported by neuronal recordings but to account for all MMN properties, complex tuning properties of neurons have to be hypothesized(May & Tiitinen, 2009).

However, the picture of a dissociation between early and late potentials as reflecting respectively automatic and conscious processing is complicated by a few findings suggesting that late potentials, including the P300 but also the N400 (Sergent et al., 2005; E K Vogel, Luck, & Shapiro, 1998) – which peaks later than the P3 and is therefore thought to be even further up the processing chain – can be elicited in unconscious conditions. First, studies reported a reduced but significant P300 (including P3b) in REM sleep (Cote & Campbell, 1999; Perrin, Garcia-Larrea, Mauguière, & Bastuji, 1999; Salisbury, Squires, Ibel, & Maloney, 1992), in response to invisible (Bernat, Shevrin, & Snodgrass, 2001; van Gaal & Lamme, 2012) and unseen stimuli (Del Cul et al., 2007; Lamy, Salti, & Bar-Haim, 2009). The increase of the P300 amplitude may therefore not systematically index conscious perception.

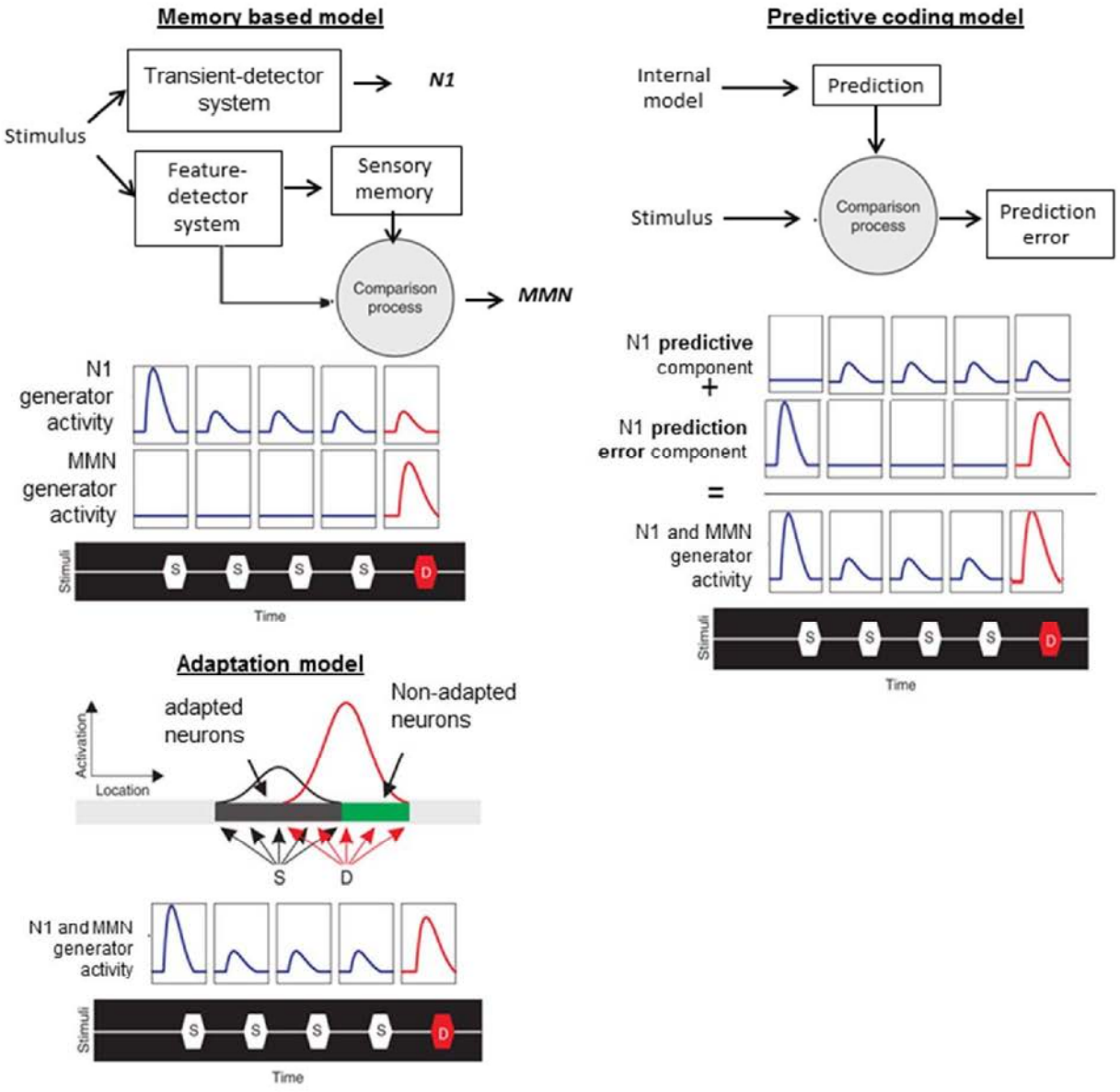


FIGURE 1.1-2: SCHEMATIC REPRESENTATION OF 3 MODELS OF MMN

In the memory-based model the MMN is generated by a comparison process distinct from the N1. The MMN is elicited by a comparison process between the incoming stimulus and a memory trace in the auditory cortex representing the repetitive aspects of the preceding auditory events, which usually lasts for a few seconds. In the adaptation model (also called "fresh afferent" model), the repetition of a stimulus leads to a synaptic habituation in the neurons encoding the standard sound. When a new sound is presented, it stimulates new, non-adapted synapses that generate a bigger response. MMN is considered the modulation of the N1 response. In the predictive coding model, the MMN is also a modulation of the N1 response. Past stimulus are used to learn an internal generative model of the stimuli that is used to predict what should come next. A prediction error signal is emitted when the stimulus is different from the predictions. The N1 response is the summation of prediction and prediction error responses. Adapted from (May & Tiitinen, 2009)

1.1.3 The "local-global" paradigm

Building on these results, Bekinschtein and colleagues (Bekinschtein et al., 2009) tried to create a protocol that would both test preserved sensory processing and assess consciousness. As the late potentials to the detection of a stimulus do not seem to be a selective enough, they tried to build a task that would not only need detection, but also consciousness-dependent processing of temporal regularities. The paradigm consisted in blocks of 125 sequences of 5 tones (Figure 1.1-3). A first level of regularity was introduced at the level of the sequence: most tones were identical (local standards), but the last one could sometimes be different (local deviant). The second level of regularity was at the (global) block level: one of the two types of sequences (xxxxx or xxxxY) was presented most of the time (70% of the trials) while the other was presented more rarely (global deviant, 30% of the trials). Crucially, in some blocs (xxxxY blocks), the local deviance could be fully expected at the global level.

In healthy attentive subjects that were asked to count deviants, the data revealed a double dissociation between the potentials elicited by the violation of the local and global regularities: *in both blocks* the violation of the local regularity elicited a response known as the mismatch negativity (MMN), which peaked between 100 and 200ms after the onset of the rare tone, even if its occurrence was predictable at the global level. The P300 component was observed in both blocks after the sequences violating the global regularity. The subjects could easily report the global rules after the experiment. In addition to the P300 response that was present in all subjects, the global deviance elicited in some subjects a response at the level of early potentials.

In this study, they author manipulated the attention of healthy subjects by instructing them to engage in mind-wandering or by giving them a challenging visual target detection task. These manipulations did not affect the MMN, whereas the P300 almost disappeared. This was consistent with the fact that few subjects in the mind-wandering group and no subject in the actively distracted group were able to report the global regularity used in the paradigm. These data suggested that conscious processing of the stimuli was necessary to elicit response to the violation of the global regularity. Most interestingly, the paradigm was also applied to patients either diagnosed as vegetative (VS), or minimally conscious (MCS) (a state intermediate between vegetative and fully conscious where the behavioral signs of consciousness are present only intermittently). The patients were instructed to pay attention to the stimuli and detect violations of the regularity. The amplitude of the response to global violation was affected in both groups compared to healthy subjects, but remained detectable in some MCS patients; whereas it was absent in all VS patients.

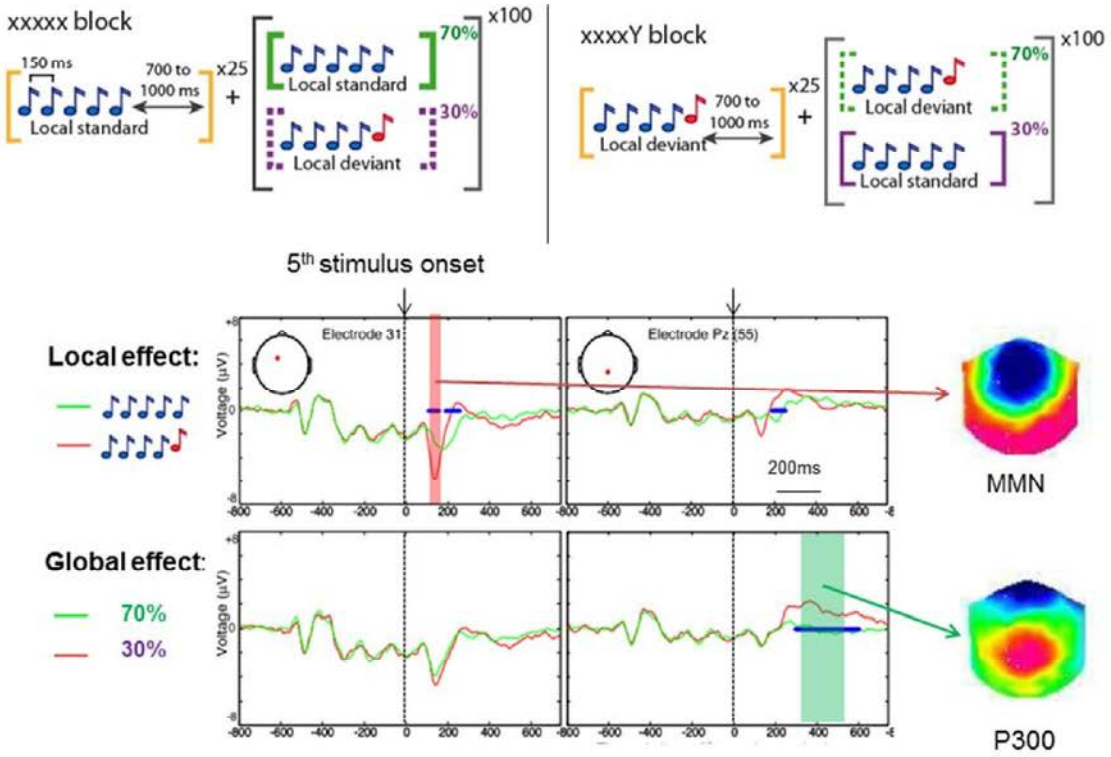


FIGURE 1.1-3: THE "LOCAL GLOBAL" PARADIGM.

Top: In this protocol, sequences of five sounds are presented to the subjects. Two types of sequences are presented in different proportion in each bloc: either the five tones are identical (xxxxx), or the first four tones are identical and the last one is different (xxxxY). In xxxxx blocks, the xxxxx sequences is presented most of the time: after 25 habituation trials where only the frequent sequence is presented, 100 sequences are presented with 70% of xxxxx and 30% of xxxxxY. In xxxxY blocks, the proportion of xxxxx and xxxxY sequences are reversed. Overall, one tone is presented most of the time (x tone) and one tone occurs rarely (Y tone); xxxxx sequences are locally standard, and xxxxY are locally deviant. In each bloc one type of sequence is frequent, and referred to as the global standard, and the other in rare and referred to as the global deviant. In xxxxx blocs, both local and global deviance correspond to the same sequences, while in xxxxY blocs, the locally deviant sequence is the global standard. Bottom: Evoked related potentials in normal subjects. Local effect: average over both types of blocks of the locally standard (xxxxx, in green) and locally deviant (xxxxY, in red) sequences. The main significant difference between ERP evoked by the two types of sequences occurs between 100 and 200ms after the onset of the fifth tone over the frontocentral electrodes and corresponds to the MMN. Global effect: average over both types of blocks of the frequent (green) and rare (red) sequences. The main significant difference between the two conditions occurs between 300 and 600ms after the onset of the last tone and corresponds to a P300 complex. The topography of the MMN and the P300 are displayed on the side. Adapted from (Bekinschtein et al., 2009)

1.1.4 Modeling challenges

This experimental protocol proved to be quite sensitive to both the state of consciousness and the attentional state of the subjects thus demonstrating the interest of using temporal regularity processing as a tool to detect consciousness in non-communicative patients. However, it opens a number of experimental and theoretical questions. First, it is unclear why the detection of the global regularity was not possible at level of processing indexed by the MMN. Indeed, MMN is thought to be elicited when "an acoustic event deviates from a memory record describing the immediate history of the sound sequence" (Winkler, 2007). What made the previous sequence too distant to be used as the context for deciding whether the last tone was deviant or not? This question is all the more intriguing given that a very close variant of the sequence of stimuli used in this experiment did not elicit a MMN to the rare tone (Elyse Sussman, Ritter, & Vaughan, 1998). Specifically, the same xxxxY sequence was repeated, but without any pause between two repetitions so that the interval between two deviants was constant. In addition, increasing the interstimulus interval (ISI) led to a recovery of the MMN response. These subtle effects are not accounted for by the current theories. Moreover the global manipulation was not completely lost for early stages of processing as a modulation of the early potentials by the global regularity seemed indeed to exist in some subjects, including one MCS patient. Understanding the limit of unconscious processing seems all the more urgent to establish meaningful tests of conscious processing.

Second, conscious processing of the stimuli seems to be crucial for the late response to exist. However, it is unclear what regularity was learned at this level, or what was the crucial computational element that allowed for the conscious detection of the rule.

In this thesis, I will therefore try to understand the neuronal mechanisms underlying automatic response to temporal regularity violations, focusing on the MMN, and derive their computational limits. We will then investigate how the properties of conscious processing might allow overcoming some of these limitations.

In the following sections, I will review the theoretical frameworks and empirical data that give us indications of *i)* the nature of the computations performed by the neuronal systems that respond to the temporal regularity violations; *ii)* the type of temporal rule that can be learned by humans and what are the characteristics of rules that would depend on conscious processing; *iii)* the neuronal properties of the systems involved.

KEY POINTS

- Determining the conscious state of non-communicative patients is an important challenge. **Neuroimaging techniques** are a powerful tool to overcome challenges that come from motor dysfunctions.
- It is crucial for clinical applications to develop protocols relying on **affordable** imaging techniques like EEG
- Only people that are conscious can access and **process consciously** information
- Theories of consciousness propose that conscious access relies on a distributed network involving frontal areas that allow the **maintenance** and the broadcasting of a **limited** amount of information through recurrent and feedback connections.
- The presence of the mismatch negativity is **independent on attention**. It can be evoked by the violation of a repetitive temporal patterns involving a wide range of features; and could thus reflect a **general property** of unconscious processing of temporal regularities
- **Late evoked potentials** seem to correlate tightly with conscious perception of targets in active tasks. They also correlate with the detection of violation of temporal regularities.

1.2 Why do we process temporal regularities?

This question is a provocative way of wondering how previous research has considered temporal regularity learning in an evolutionary point of view. In other words, this section tries to capture in what ways the capacity to learn temporal regularities can confer an evolutionary advantage to an organism. Looking for an answer to this interrogation is not merely a philosophical quest: it corresponds to the first conceptual level of Marr's hierarchy (Marr, 1982), namely the computational level. Indeed, understanding what problem the brain is trying to solve is the first step in explaining its function. It may also help to answer the question: what is the function of the potentials evoked by violations of temporal regularities like the P300 and the MMN?

1.2.1 Predictive coding

A first answer comes from the increasingly popular framework of predictive coding. This approach builds on the observation that most of the information in sensory inputs is *redundant*: natural sensory inputs are produced by objects in the physical world that tend to be coherent spatially (they extend in space) and temporally (they last in time). As a result, adjacent pixels on an image are likely to be similar; an edge is likely to follow a continuous and smooth trajectory; an object that was present the moment before is likely to last. This translates into correlations between pixels of an image (Dong & Atick, 1995) or sounds (Lewicki, 2002) over both time and

"space" (between pixels or between frequencies). A direct representation of the raw image or sound by the activity of an array of sensory receptors would thus be very inefficient. It has long been suggested based on information theoretic considerations (Attneave, 1954). A more *efficient* way of encoding information would be (1) compress the information, eliminating the redundancies by trying to infer the hidden states of the environment that are responsible for them, and then (2) only encode the sensory information that violates the regularities that can be predicted based on the inferred model (R. P. Rao & Ballard, 1999), i.e. the prediction errors.

Predictive coding proposes that the perceptual systems try to learn an internal model of the environment and use this model to actively predict the incoming signals. A basic assumption is that the world produces inputs on the sensory receptors in a way that can be described by a dynamical generative model. These model can typically be expressed by a system of equations of the form (Friston & Kiebel, 2009a):

$$\begin{aligned} y &= g(x, v, \theta) + n \\ \dot{x} &= f(x, v, \theta) + w \end{aligned} \tag{1}$$

where f and g are potentially nonlinear functions parameterized by θ. g describes how the *causes* v, which are invariants aspect of the world such as objects (Shipp, Adams, & Friston, 2013b), influence the input data y. The states x represent "hidden states" that describe how causes interact with each other and endow the model with memory. f describes the *temporal dynamics* of the these states. n and w are Gaussian noise processes representing observational noise.

The optimal way the brain can encode information to minimize prediction error, is *i)* to *learn* the model of the world that will lead to minimal prediction errors on average. This is often considered to be the actual generative model of the input, and *ii)* to be able to *infer* the appropriate causes and states at a given time to best explain the input.

This idea rejoins a more classical idea in philosophy of perception which states that perception is not merely the reflect of sensory input but an *hypothesis about their causes* in the external world (Helmholtz, 1860). Moreover, sensations contain insufficient information to infer what caused them without additional prior: for example, the retina captures only a 2D projection of a 3D world, which implies that the source space is of higher dimensionality than the sensor space, making the inference of causes an ill-posed problem. Locke (Locke, 1690) was already

proposing that "our minds should often change the idea of its sensation into that of its judgment, and make one serve only to excite the other": in other terms, we should have prior expectations about the causes of our inputs and only change these beliefs if the inputs contain inconsistent information.

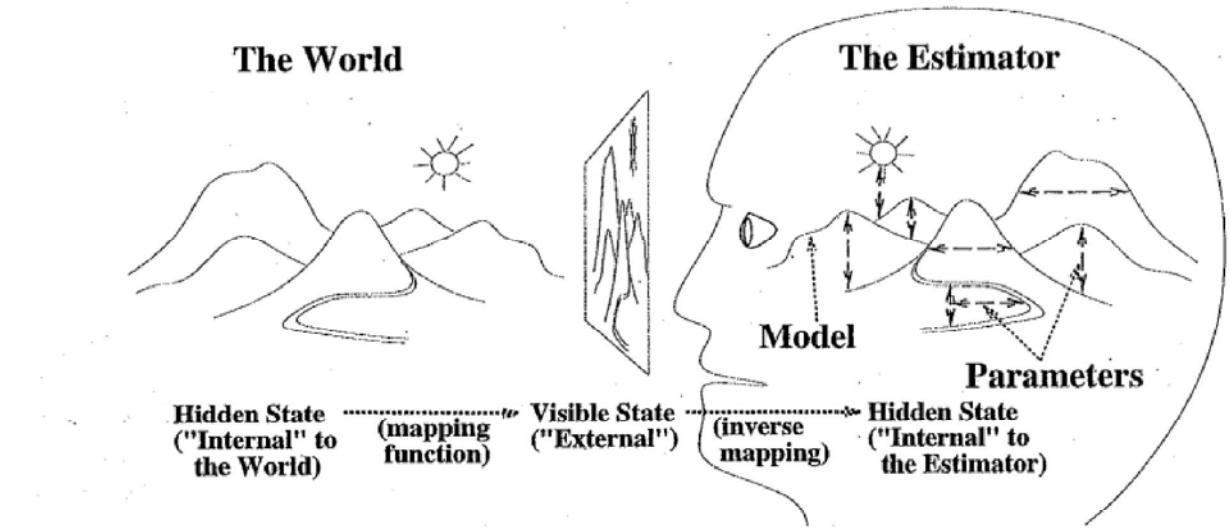


FIGURE 1.2-1: INTERNAL MODELS.

The problem faced by an organism relying on an internal model of its environment. The underlying goal is to optimally estimate, at each time instant, the hidden state of the environment given only the sensory measurements. This can be done by building an internal model of the generation of the sensory input that approximate as best as possible the true model. From (R. Rao & Sejnowski, 2002)

More realistically, generative models are considered herarchical: features at larger temporal or spatial scale (the song, the object) determine how lower scale features (the pitch of a particular tone, the edge of the object) are going to evolve. This type of model corrspond to a generalization of equation (1).

$$
\begin{aligned}
y &= g\left(x^{(1)}, v^{(1)}, \theta^{(1)}\right) + n^{(1)} \\
\dot{x}^{(1)} &= f\left(x^{(1)}, v^{(1)}, \theta^{(1)}\right) + w^{(1)} \\
&\vdots \\
v^{(i-1)} &= g\left(x^{(i)}, v^{(i)}, \theta^{(i)}\right) + n^{(i)} \\
\dot{x}^{(i)} &= f\left(x^{(i)}, v^{(i)}, \theta^{(i)}\right) + w^{(i)} \\
&\vdots \\
v^{(m)} &= \eta + n^{(m+1)}
\end{aligned}
\tag{2}
$$

with the $g^{(i)} = g\left(x^{(i)}, v^{(i)}, \theta^{(i)}\right)$ functions linking the i^{eme} and (i-1)^{eme} hierarchical levels of causes, and the $f^{(i)} = f\left(x^{(i)}, v^{(i)}, \theta^{(i)}\right)$ functions describe the dynamics of the hidden states *within* a hierarchical level.

From this generative model it is possible to derive a likelihood function $p(y|x,m)$ which specifies the likelihood of some data given the causes and the model m. Inverse this model to determine which are the most likely causes of the input is a good way to minimize prediction errors. This can be done using variational Bayes, which is based on Bayes rule: $p(x|y,m) = \frac{p(y|x,m)p(x,m)}{p(y,m)}$ that relates the likelihood of the causes given the data $p(x|y,m)$ to the internal generative model that specifies the likelihood of the data given the causes. Because the generative model is hierarchical, Bayesian inference is itself a hierarchical process.

Predictive coding proposes that the brain learns a model of the world, and then inverts it to maximize coding efficiency and "explain away" predictable information. But is this coding scheme consistent with neuronal data?

First, a key architecture principle of the brain is its hierarchical organization (Felleman & Van Essen, 1991), particularly clear in the early visual areas. The notion of hierarchy in the cortex relies on the distinction between forward and backward connections. Forward connections arise from pyramidal cells in the supragranular layers (layers 2-3) and terminate in spiny stellate cell of layer 4 of the higher cortical area. Feedback connections arise mainly from pyramidal cells in the infragranular layers (layers 5-6) and target both infra and supragranualar layers of the hierarchically lower area. The idea of a hierarchy that arises from these anatomical definitions is also present in functional aspects of neuronal response which show an enlargement of spatial and temporal (Gauthier, Eger, Hesselmann, Giraud, & Kleinschmidt, 2012) receptive fields from the lower areas to the higher areas and representations that go from simple to more complex and abstract (DiCarlo & Cox, 2007; Hubel & Wiesel, 1968). This general architecture allows for mapping of a hierarchical generative model onto the cortical substrate, with feedforward connections sending remaining prediction error to the higher areas, and backward connections mediating the transmission of predictions to lower areas.

Second, predictive coding has been successful at explaining neuronal responses at multiple levels of the brain's hierarchy, from the retina (Hosoya, Baccus, & Meister, 2005) to V1 (R. P. Rao & Ballard, 1999) and could be implicated in extra receptive fields or temporal prediction effects described in other areas including MT or IT cortex (Huang & Rao, 2011; Jehee,

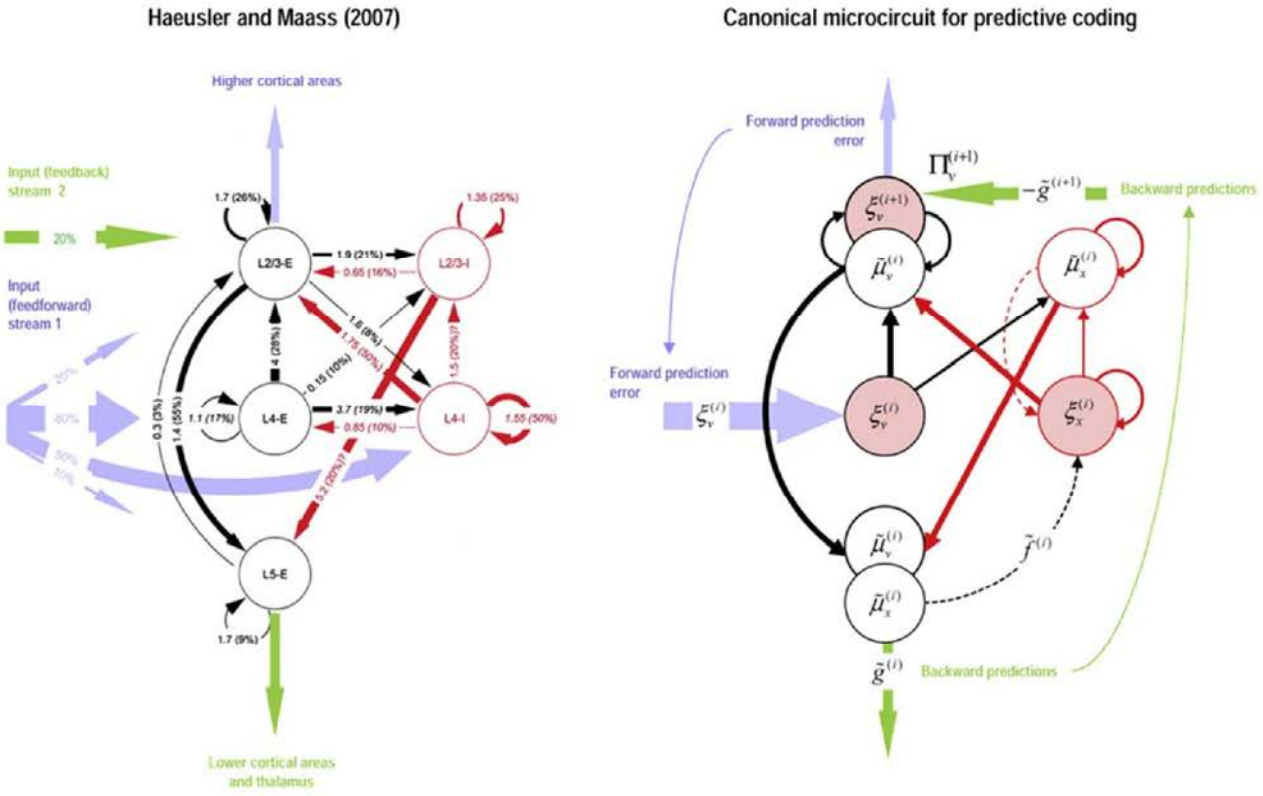
Rothkopf, Beck, & Ballard, 2006; Meyer & Olson, 2011). It was used to explain extra-receptive field effects in vision. For example, the response of a V1 neuron sensitive to oriented bars in its receptive field sees its response suppressed if the bar extends outside of the receptive field. Hierarchical predictive coding explains this intriguing propriety by showing that information from neighboring receptive fields inform the higher level area, allowing it to better predict the content of the first receptive field (R. P. Rao & Ballard, 1999). At the molecular level, the modulation of the neuronal response by feedback predictions appears to be mediated by NMDA receptors (Self, Kooijmans, Super, Lamme, & Roelfsema, 2012). Experimental manipulation of the statistics of the input to the retina were shown to result in modifications of the receptive fields of retinal ganglion cells (Hosoya et al., 2005) that improved encoding efficiency as predicted by the predictive coding framework. Interestingly, the presentation of a given context (flickering uniform stimuli or checkerboards) for 13s was sufficient to drive adaptation of the receptive fields to the spatial statistics of the input: cells presented with flickering checkerboards became *less* sensitive to checkerboard by a factor 0.57, and *more* sensitive to uniform stimuli by a factor 1.4. Manipulation of the temporal correlation modulated the temporal filter of the cells, without accounting fully for the statistics of the input.

Most of predictive coding research efforts based on neuronal data have concentrated on perceptual inference on stationary inputs, i.e. on the inference about causes $v^{(i)}$ rather than inference about states. A notable exception is the work of Friston and Kiebel (2009b) that showed how the variational Bayes framework could allow predictive coding of very complex hierarchical temporal patterns reproducing a birdsong-like sequence of sounds, given that the generative model of the song was already learned. They showed in particular that the omission of the end of the temporal pattern in the input elicited an omission response corresponding to the prediction error of the predicted but absent sounds. Friston (Friston, 2005) proposed that MMN was the correlate of such temporal predictive coding, where repetition of a stimulus at regular interval would lead to a modification of the internal model of the inputs, just like habituation to the checkerboard lead to a modification of the receptive field of the retinal ganglion cells of the retina in a few seconds. This would lead to an attenuation of the response which would be revealed when a deviant tone is presented. Note that just based on the oddball paradigm data, both a hierarchical implementation actively predicting a repetition and synaptic habituation are consistent with the predictive coding principle: the response to repeated inputs is attenuated. However, a response to expected sounds that are omitted are observed at the latency of the MMN (Bendixen, Schröger, & Winkler, 2009b; Raij, McEvoy, Mäkelä, & Hari, 1997; Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997) as predicted by predictive coding and have been

cited as a decisive argument to rule out the simple synaptic habituation hypothesis (but see May & Tiitinen, 2010).

FIGURE 1.2-2: CANONICAL CORTICAL MICROCIRCUIT AND PREDICTIVE CODING

(Left) Cortico-cortical connections in the canonical microcircuit. Data from Haeusler and Maass (2007). Layer 4 neurons receive the incoming feedforward input. Excitatory connections go from layer4 to supragrannular layers, to infragrannular layers. (Right) Proposition of correspondance between the neuronal populations from the canonical microcircuit and the variables of a predictive coding algorithm based on variational Bayes. μ and ξ represent respectively predictions and prediction errors about states x and causes v. From (Bastos et al., 2012).



Given the success of the predictive coding framework in explaining the neuronal data, a pressing question is to understand how the different computational aspects of the algorithm can be implemented in neuronal circuitry. Diverse mapping of the crucial units – namely units encoding the predicted states and causes, and units encoding prediction errors about states and causes – onto the cortical layers and the canonical cortical microcircuit have been proposed, along with correspondences between anatomical connectivity and the flow of information predicted by Bayesian inference (Bastos et al., 2012; Shipp et al., 2013b; Spratling, 2012). In these models, layer 4 units are thought to represent prediction error. Supragranular layers would contain both predictive and prediction error units. Infragranular layers would only contain predictive unit and serve as a relay for descending predictions. Inference about causes tend to

involve both intra and inter area connections, whereas the computations concerning states are mainly the result of local connectivity, consistently with the generative model (2) showing that temporal evolution of hidden states at the i^{th} level depends only on the parameter of this level.

We know very little of the neuronal basis of model learning. (Hosoya et al., 2005) proposed a spatial predictive scheme based on anti-hebbian plasticity, so that units representing similar information tend to inhibit each other. Training of spatiotemporal filter based on natural input involved the computation of optimal filter based on minimization of arithmetic function and have little to say about the way these optimizations are computed in the brain (R. P. Rao & Ballard, 1999).

Moreover, predictive coding is a generic framework that argues more in favor of a "universal" computing unit that can be replicated across cortex at every level of the hierarchy. If MMN is a potential characteristic of the computations of a predictive unit, is should not be specific of auditory cortex. This is indeed the case: an equivalent of the auditory MMN can be found in other sensory modalities including visual (Pazo-alvarez, Cadaveira, & Amenedo, 2003; Tales, Newton, Troscianko, & Butler, 1999), olfactive (Krauel, Schott, Sojka, Pause, & Ferstl, 1999; Pause & Krauel, 2000) and somatosensory (Kekoni et al., 1997; Shinozaki, Yabe, Sutoh, Hiruma, & Kaneko, 1998) modalities. In the auditory modality itself, MMN responses are observed to a wide range of features like sound frequency (Bekinschtein et al., 2009; Näätänen et al., 1978), intensity (Javitt, Steinschneider, Schroeder, & Arezzo, 1996), spatial location (Deouell, Parnes, Pickard, & Knight, 2006), duration (Näätänen, Paavilainen, & Reinikainen, 1989)but also conjunctions of features (Gomes, Bernstein, Ritter, Vaughan, & Miller, 1997) and more complex features like change in vowel (Dehaene-Lambertz, 1997). In auditory cortex, the sources for duration and frequency deviance were found to be separable (Sysoeva, Takegata, & Näätänen, 2006), which suggests that separate modules detect deviance regarding the different features of the stimulus. Overall these data support the idea of a broad computational significance of MMN as a shared mechanism across the respective sensory hierarchies responsive to unpredicted stimuli. However, the generality of this approach brings very little insight about the potential consciousness-dependent computations reviewed earlier, that would depend on a central workspace with specific connectivity patterns. We have to look for clues about the role of conscious access in temporal regularity processing in other approaches.

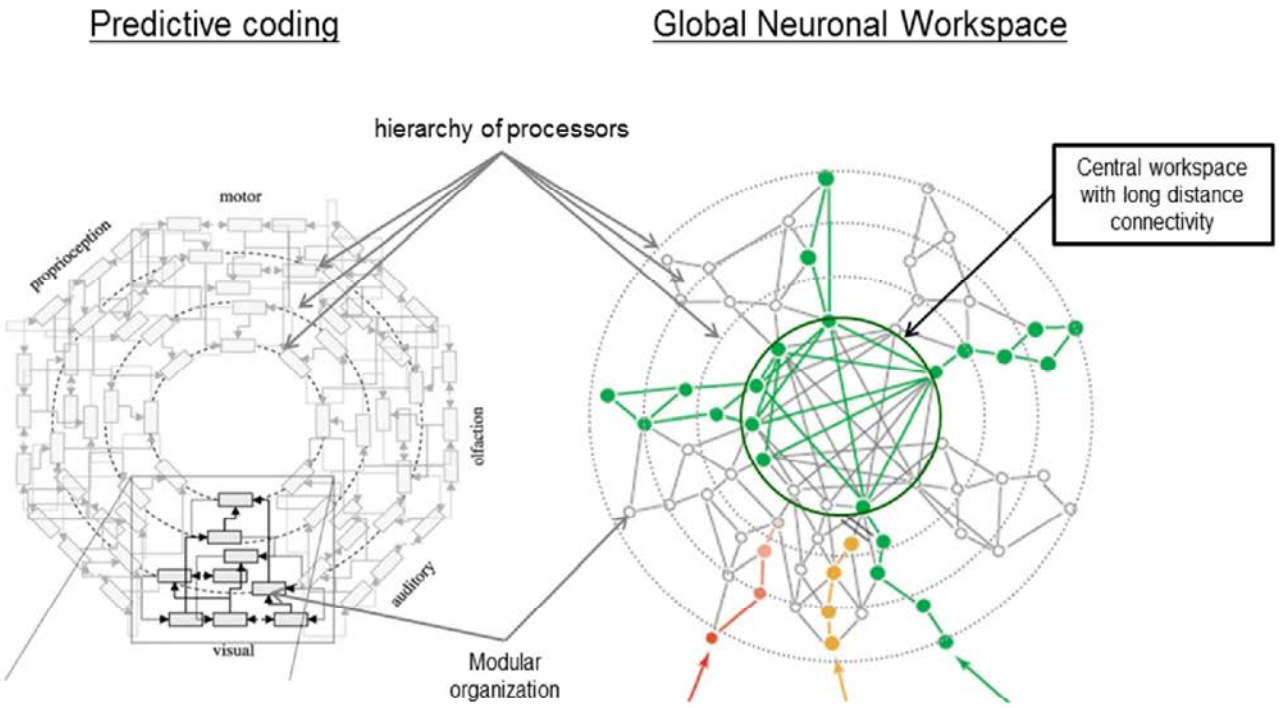


FIGURE 1.2-3: HIERARCHICAL PREDICTIVE CODING AND GLOBAL NEURONAL WORKSPACE ARCHITECTURES

Comparison of the architecture of the two frameworks. Both models are based on a hierarchy of modular processors. Only the global neuronal workspace theory emphsizes the importance of the long range connectivity between higher level areas maintaining a stable broadcasted representation. Adapted from Friston (2005) and (Stanislas Dehaene et al., 2006).

KEY POINTS

- Predictive coding is a general coding scheme that aims for efficient coding of information
- Efficient coding can be achieved by learning a model corresponding to the generative model of the world
- This generative model involves causes that represent invariant objects that produce regularities, and hidden states, that describe their temporal dynamics
- Predictive models are hierarchically organized and inference rely on hierarchical interactions that follow the feedforward and feedback connections between cortical areas
- Prediction error units are thought to be found in layer 4 and layers 2-3, predictive units should be in supra and infragranular layers.
- Increased response to unexpected events is consistent with predictive coding.
- Little is known about neuronal substrates of model learning.

1.2.2 Behavioral consequences of temporal regularity learning

The main argument of the predictive framework in favor of regularity learning is efficiency of encoding. However, an organism that has efficient encoding of information but inappropriate behavior is unlikely to survive for long. There are multiple behavioral responses which reveal that animal and humans exploit temporal regularities in natural inputs to inform their behavioral response.

1.2.2.1 Reward anticipation and decision making

One of the first evidence that temporal regularities are indeed processed by the brain, and result in adaptive behavioral responses comes from conditioning experiments. Conditioning paradigms can be schematically split into two traditions that highlight two behavioral functions of temporal predictions.

The first one was initiated by Pavlov (Pavlov, 1927) who studied how the systematic pairing of a neutral stimulus with a rewarding or punishing stimulus, could lead to the acquisition of a new behavioral response to the occurrence of the neutral stimulus. Specifically, he first showed that the repeated temporal pairing of the ringing of a bell with food ended up eliciting salivation in response to the bell alone. Pavlov called this response a "conditional reflex" because salivation is a reflex response. In the case of the bell, the reflex response is not hard "hard coded" (i.e. it was not learned on evolutionary time scales), but learned on shorter time scales depending on the environmental *conditions*. In conditioning, the stimulus producing the automatic response (the food) is called an Unconditioned Stimulus (US) whereas the initially neutral stimulus is called the conditioned stimulus (CS). The responses elicited by these two stimuli are called respectively the unconditioned (UR) and conditioned (CR) responses. Crucially, the CR does not appear as an immediate response to the CS. If the US was always presented at a fixed delay after the onset of the CS, then the CR tends to happen at the time where the US was *expected* to happen : when an animal is conditioned to expect an air puff eliciting an eye blink when a tone is presented, it does not only blink to the tone, it blinks at the time the air puffs would normally occur (Kehoe, Graham-Clarke, & Schreurs, 1989), making this association extremely adaptive at the behavioral level. Interestingly, the association between CS and US is not systematic. In the blocking paradigm, the animal goes through two phases of learning: first the US is pared in a delay conditioning-like protocol with a CS – say a tone, that we will call *CS1* – until the association is learned. Then in a second phase, *CS1* is presented at the same time as a second CS – say a light, *CS2* – while still paired with the US. Even though the temporal relation is similar between the

two CS and the US, the association between *CS2* and the US is never learned. This result can be interpreted as evidence that a new association cannot be learned if the US is already predicted.

A large range of effects in classical conditioning can actually be accounted for by the Rescorla-Wagner model (Rescorla & Wagner, 1972). It specifies the dynamics of the evolution of the strength V_X of an association between a stimulus X that has a saliency α_X and a US.

$$\Delta V_X = \alpha_X \beta (\lambda - V_{tot})$$

β represents a free parameter corresponding to the association value of the US, that describes how "meaningful" a US is. λ is the maximum conditioning value for the US, usually set to 1 when it is present, 0 otherwise. V_{tot} is the sum of all the association strength between the US and other stimuli. If two CS (X1 and X2) are paired with the US. $V_{tot} = V_{X1} + V_{X2}$. This model successfully accounts for a large number of properties of classical conditioning, including the blocking paradigm, as it relies on a *prediction error* term $(\lambda - V_{tot})$ that predicts that no learning can happen if the US is already completely predicted.

The second type of conditioning, called operand or instrumental conditioning consists in associating actions with rewards or punishments. It which can be traced back to Thorndike's "law of effects" (Thorndike, 1911) which was expressed in this way: "Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will be more firmly connected with the situation, so that, when it recurs, they will be more likely to recur". Following this law, it is widely believed that animals and humans make decisions in order to maximize rewards and minimize losses or punishments. To this end, the animal or the human has to be able to infer which outcome can be expected from each possible action. The field of reinforcement learning developed algorithmic solutions to that problem (Doya, 1999; R.S. Sutton & Barto, 1998a). The simplest solution is called model-free estimation. It is considered to be a retrospective method, because it consists essentially in storing for each possible state s and action a a value $Q(a, s)$ that reflects the past outcomes encountered when the action was chosen in the past. This value can be interpreted as a prediction of the expected outcome. This value is updated each time the action is chosen according to the temporal difference (TD learning) rule:

$$Q(a, s) \leftarrow Q(a, s) + \eta\delta$$

where η is the learning rate, and

$$\delta = r(t) + \gamma\, Q(a', s') - Q(a, s)$$

is the reward prediction error, with a' and $s'$are respectively the next action and the next state and $0 < \gamma < 1$ a time discounding factor. This update algorithm is closely related to the Rescorla-Wagner model, as leaning is similarly dependent on a learning rate (η for reinforcement learning, $\alpha_X\beta$ for classical conditioning) and a prediction error term (respectively δ and $(\lambda - V_{tot})$). A positive error signal can be viewed as a "good" surprise, meaning that the outcome was better than predicted. Crucially, the correlate of the reward prediction error signal postulated by TD-learning was reported in the dopamine neurons in midbrain (Schultz, Dayan, & Montague, 1997) projecting to the dorsolateral striatum, the nucleus accumbens and possibly cortex, in a classical conditioning paradigm. The response to a predictable reward shifts from the time of the reward at the beginning of the protocol to the predictive cue (CS) at the end of reinforcement (Figure 1.2-4). The existence of a predictive inhibitory signal can be observed through the decrease in firing rate observed when the reward is omitted. This finding was largely reproduced since in the context of operand conditioning (Kobayashi & Schultz, 2008). It is interesting to observe that classical conditioning – in which the relation between the stimulus and the reward does not depend on an action – and instrumental conditioning – in which it does – rely on similar algorithms that use prediction errors to adapt a predictive model.

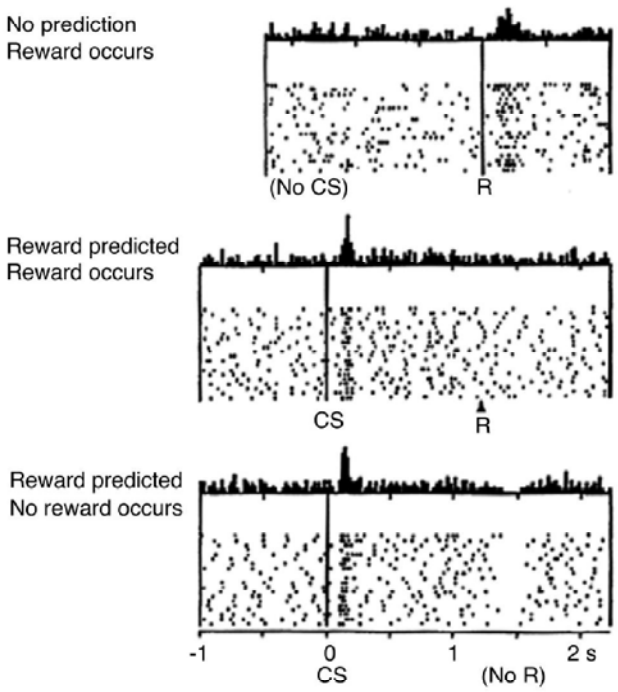


FIGURE 1.2-4: RESPONSE OF DOPAMINE NEURONS TO REWARD AND PREDICTED REWARD

Recording from a dopaminergic neuron. Each line of the raster plot represents a trial, the histogram represents the average firing rate of the neuron. (top) when no stimulus predicts the reward, the neuron respond right after reward delivery. (middle) After classical conditioning, the neuron responds after the predictive stimulus and does not respond any more to the reward.

(bottom) If the reward is omitted, the firing rate of the neuron falls below its baseline at the timing of the expected reward. From (Schultz et al., 1997)

But conditioning relies on rewarding or punishing stimuli in both of the cases previously described. Is it reasonable to think that the same kind of algorithm could also be used in the absence of a reward signal? Could we imagine learning the internal model used in perceptual inference using a similar process?

Conditioning experiments give us reasons to think so. First, conditioning can be observed even though the association between a stimulus and the reward has never been experienced is some specific cases. In the protocol called second order conditioning, an animal is first trained to associate two CS that have a reproducible temporal relation – say a sound paired with a light. Then one of the CS (for example the light) is associated to a US. In the third phase, the sound is presented alone. Interestingly, the sound elicits a conditioned response, as if it had been paired with the US (Holland & Rescorla, 1975). This result suggests that the relation between the two CS was learned in the absence of a reward or punishment. A similar phenomenon can be observed in operand conditioning and is known as latent learning (Blodgett, 1929). Blodgett trained two groups of rats to find food in a maze. The first group was trained with food from the beginning, but the second group was first left in the maze without food a few hours a day for a few days, where they could explore the maze without any reward. He observed that when food was introduced in the maze, the learning curve from the second group was much steeper than the learning curve of the first group, as if prior experience was facilitating learning. Tolman (Tolman, 1948) proposed that the animals could, in the absence of reward, build cognitive maps that represent a model of the world. This model of the world can then be used to reach rewards faster once they have been introduced.

Latent learning seems to share the properties of classical CS-US association, like the capacity to learn temporally precise information. In second order conditioning, the timing of the relation can be used in the second order transfer. In an experiment (Arcediano, Escobar, & Miller, 2005), subjects integrated the association US→CS1 with a second CS2→CS1. By using the CS1 element found in both association, and by integrating the respective timing of the associations, subjects formed a new predictive association between CS2 and US, only if the respective timing of the associations predicted that CS2 would precede the US (i.e. be predictive). Both reward prediction error and state prediction error (occurrence of a surprising CS)have been shown to coexist in humans (Gläscher, Daw, Dayan, & O'Doherty, 2010).

The Rescorla-Wagner model and the TD-learning algorithm both fail to explain these transfer behaviors, as they rely on direct experience of the association for learning. A second class of models has been developed to account for these behaviors called model based evaluation. Model based methods are thought of as prospective, relying on predicting future rewards given possible future scenarios that are imagined based on a model of the transitions between future possible states and the rewards associated with them. Latent learning would consist in learning this model of transitions between states. Model based methods allow much more flexibility than model free methods in changing environments, but they are also much more computationally intensive because they rely on a simulation of all the possible future scenarios whereas model free requires only the retrieval of a stored value. How the brain arbitrates between these two systems is still debated (Daw, Niv, & Dayan, 2005; Dayan & Niv, 2008; Otto, Gershman, Markman, & Daw, 2013). However one of the factors influencing the evaluation method used appears to be working memory load: a concurrent demanding secondary task biases choices towards model free estimates (Otto et al., 2013), suggesting that the computation of values based on an internal model depends on central resources with limited capacity.

KEY POINTS

- Classical conditioning provides evidence that the temporal structure of events is exploited to adapt the timing of automatic responses in reaction to aversive and rewarding events.
- Instrumental conditioning suggests that the temporal structure of rewards is exploited to inform decision making in order to choose actions that lead to the best outcomes
- Dopaminergic neurons in the midbrain respond to predictable rewards in a way consistent with TD-learning
- The same kind of algorithm could be used to predict transitions between events even in the absence of rewards, i.e. to learn a model of the world.

1.2.2.2 Motor control optimization

In addition to exploiting regularities in the sensory inputs, people seem to be able to capture regularities in their motor behavior. In the SRT task, a target such as a dot appears in one of several possible locations on a computer display and the participant presses as fast as possible a response key assigned to that location. Instead of appearing at random across a series of trials, however, the target follows a predictable or partially predictable sequence of locations. Learning is measured chronometrically by interleaving random sequences between blocs of regular sequences; an increase in reaction time (RT) on the random sequences is evidence that participants did not simply improve their general motor skill but also learned something specific

about the temporal structure of the training sequence and were using their knowledge to anticipate the target location on each trial, thus achieving rapid RTs learning (Nissen & Bullemer, 1987). This decrease in RT is also accompanied by a reduction of the error rate (Haider, Eichler, & Lange, 2011).

Decrease in RT is an evidence for prediction allowing anticipation, but does not prove that the learning occurs at the motor level. How can we be sure that this effect is not purely the result of perceptual learning? In an experiment (Willingham, 1999), subjects where trained on a SRT task for a few blocks leading to improvement of reaction times. The stimulus-response mapping was then changed so that one group of subjects pushed the same sequence of keys but saw new stimuli, whereas another group pushed a different sequence of keys but saw the same stimuli. Transfer to the new mapping was shown only if the motor sequence was kept constant, not the perceptual sequence, showing that the learning is indeed motor.

Motor learning can occur without any consciousness that it happened, or that there was a regularity (Destrebecqz & Cleeremans, 2001). On some cases, an explicit awareness of the knowledge can emerge. The emergence of a knowledge reportable verbally seems to be facilitated by violations of the learned regularity (Haider & Frensch, 2005, 2009; Rünger & Frensch, 2008). The behavioral effects of awareness of the regularity on behavior include additional facilitation of RT that appear in a step like fashion (sudden drop of RT) and the capacity to resist perceptual interferences in Stroop-like task (Haider et al., 2011).

KEY POINTS

- Predictable motor sequence are produced faster and more accurately that unexpected sequences
- In motor domain learning can be independent of consciousness

1.2.2.3 Orienting of attention towards new events

Consistently with the idea that prediction errors can facilitate conscious access, violations of temporal regularities have been shown to attract attention since Pavlov's work on conditioning. The father of conditioning described the orienting response as follow:

"I call it the "What-is-it?" reflex. It is this reflex which brings about the immediate response in man and animals to the slightest changes in the world around them, so that they immediately orientate their appropriate receptor-organ in accordance with the perceptible quality in the agent bringing about the change, making full investigation of it.

The biological significance of this reflex is obvious. If the animal were not provided with such a reflex its life would hang at every moment by a thread. In man this reflex has been greatly developed in its highest form by inquisitiveness-the parent of that scientific method through which we hope one day to come to a true orientation in knowledge of the world around us" (Pavlov, 1927)

This description shows the behavioral importance of deviance detection: unexpected events bring uncertainty about possible danger or rewards. Learning temporal regularities allows the disengagement of attention from innocuous predictable events and the signaling of elements that are worth investigating. Sokolov's experiments (Sokolov, 1963) showed that orienting response was not elicited by salient stimuli, but by any stimuli that deviated from the standard one. For example, if habituated with a white noise, an animal would show orienting response to increase in the noise loudness, but also to attenuation of the sound or even omission. Changes in the duration or in the envelope of the noise would be detected and attended as well. The orienting reflex has been shown to habituate after after a few repetitions

Electrophysiological investigations have proposed that the correlate of the orienting response in ERP is the P3a, also called the "novelty P300". As reviewed previously, the P3a potential is typically elicited by rare stimuli, even when they are not task-relevant (N. Squires et al., 1975). The novelty P3 responses habituate across successive presentations of novel items, indicating that as these stimuli become more predictable, the magnitude of the response wanes rapidly (R T Knight, 1984). It is not tied to any particular modality — similar novelty P3 responses have been observed for novel visual, auditory and somatosensory events (R T Knight, 1984; R. Knight, 1996; R.T. Knight, 1997). It has been associated with the bottom up orienting of attention, and is thought to reflect more attentional processes than the computations leading to the detection of stimulus deviance *per se*.

The automatic attraction of attention by contextually novel stimuli can be interpreted as an intrinsic motivation to reduce predictive uncertainty about the stimulus. Consistently with this idea, it is possible to condition an animal to prefer a cage rather than another by systematically introducing novel items in it (Bevins & Bardo, 1999; Bevins, 2001). Additionally, if given a task where their choices have no impact on the outcome but can lead to early information about the reward, monkeys prefer the situation where the cues about the amount of reward are the most informative (Bromberg-Martin & Hikosaka, 2009). Finally, animal avoid actively situations that are ambiguous, where uncertainty cannot be reduced, even if they lead objectively to better outcome (Hayden, 2010). Together, these results suggest that an intrinsic motivation of the brain

leads to the deployment of executive and attentional resources in order to reduce prediction error over next events. Interestingly, place conditioning by novelty suggests that rewarding signals could come not from the minimization of prediction error but from the maximization of its diminution.

KEY POINTS

- Attention is attracted in a bottom up fashion by "novel events"
- The orienting response to novel event correlates with the P3a potential, which can be interpreted as a prediction error response
- The reflexive allocation of attentional resources to uncertain situations suggests that the reduction of uncertainty and the minimization of prediction error are intrinsic goals of attention-dependent processes

1.2.3 Free-energy principle

An integrative framework reconciles the predictive coding framework and the behavioral manifestations of regularity processing. Friston and collaborators (Feldman & Friston, 2010; Friston & Kiebel, 2009b; Friston, Kilner, & Harrison, 2006; Friston & Stephan, 2007; Friston, 2005; Kiebel, Kriegstein, Daunizeau, & Friston, 2009; Shipp et al., 2013b) propose to consider not only perception, but the action-perception cycle as a whole. They observes that living organisms are highly structured beings and seem to escape the second law of thermodynamics. This principle dictates that entropy should increase, i.e. that structure should disorganize with time. The only way to escape the second principle to maintain an homeostasis far from equilibrium is to associate correct inference about the environment through perception and appropriate actions. As an analogy Friston (Friston et al., 2006) propose to examine the destiny of a snowflake falling through the sky. A normal snowflake is a self-organized dissipative system that is unable to act on its environment. It cannot avoid falling and will necessarily meet a phase transition and loose its physical integrity. If we now imagine that the snowflake has wings and sensory input allowing it to judge its altitude; it can act on its environment to regulate its altitude and the temperature of its environment and could in theory avoid phase transition indefinitely. For that, it has to *restrict* itself to a domain of parameter space that is far from the phase-boundary. Friston argues further that evolution has *necessarily* selected organism that had developed this capacity to remain into a bounded area of the parameter space. In other words, the entropy of the sensory states must be limited. He proposes that a successful strategy to achieve this goal is the minimization of a quantity called *free-energy* which is derived from statistical thermodynamics and represents a superior bound on entropy. The minimization of free-energy

corresponds to the minimization of prediction errors through optimal inference of the causes of the input in perception, and through actions that lead to expected inputs.

Using this framework, it was possible to propose accounts of phenomenon as diverse as perceptual inference, decision making, attentional effects or the absence of a granular layer in motor cortex (Shipp, Adams, & Friston, 2013a).

However, like predictive coding, this framework gives little information about potential qualitative dissociations between conscious and unconscious processing.

KEY POINTS

- Temporal regularity learning can be found in multiple domains of behavior, cognition and neuronal coding
- In all of these domains, theoretical models relying on predictive processes have been proposed
- Predictions of expected future events have multiple advantageous roles:
    - reduction of the redundancy of encoded information leading to more efficient neuronal coding
    - better efficacy of behavioral responses (RT and error rates),
    - capacity of making decisions that lead to optimal outcomes
- Predictions errors are useful for:
    - updating of the perceptual inference about the causes of an sensory input,
    - correcting of the internal predictive model,
    - attracting attention towards unpredicted events.
- The mismatch negativity can be interpreted as a prediction error response.

1.3 What kind of regularities can we learn?

The first section of this thesis showed that unconscious perception was thought to rely on a multitude of parallel processors operating locally, while conscious perception was characterized by an ignition of a unique central space that maintains and broadcasts its content. The previous section highlighted the importance of the predictive coding framework in understanding the computational significance of neuronal and behavioral responses to predictable and unexpected stimuli. The elegance of the predictive coding framework resides in its unifying power. However, these theories do not address the question of consciousness and provide little insight into qualitative discontinuities in temporal regularity processing along the cortical hierarchy. Yet, one goal of this thesis is to better understand the computational capacities and limits of unconscious

temporal regularity processing. In other words I aim at understanding what type of property makes a repeated temporal pattern impossible to learn if its constitutive elements are not granted conscious access.

In this section we will look at frameworks and experimental data that have attempted to propose qualitative distinctions between types of sequences on the basis of computational demands and evaluate how relevant these frameworks are to our questions.

1.3.1 Chomsky hierarchy: a framework for the classification of sequential regularities

The most systematic classification of sequential regularities is probably Chomsky hierarchy. Established to clarify the level of complexity of a language, it relies on the computational demands that are necessary to generate the sequences that belong to a given language.

1.3.1.1 Formal language theory (FLT)

Formal language theory proposes a "purified" approach to the study of grammar in natural languages; purified in the sense that a number of characteristic of natural language that are not part of syntax have be removed. FLT focuses on strings of arbitrary symbols. These symbols have no meaning and no relations between them. There is no notion of words and no clues about the underlying structure of the string – as punctuation of prosody would do in natural languages – only series of symbols.

In FLT, a language is defined as a set, potentially infinite, of possible series. A grammar is defined as a procedure that allows deciding whether a series of characters belongs to the language or not, which makes it a particular case of temporal regularity. Grammars can take multiple forms, from the enumeration of all possible series of symbols, to the enunciation of a compact procedure to generate all possible sequences (generative grammar) or to verify whether a sequence is possible. Importantly, there is no notion of probability in FLT: a string does or does not belong to the language.

Generative grammars are formalized as a series of transformations that allow to go from a initial state to a possible string by applying some rules that transform symbols or series of symbols into other symbols. These rules can be seen as the generative model of the language. Two type of symbols are used : terminal symbols, that constitute the final string (noted with lower case letters), and non-terminal symbols, that cannot be present in the final strings and have to be rewritten using one of the grammatical rule before a possible string can be obtained. A

generative grammar is defined as a set of initial state S, terminal Σ and non-terminal N symbols, and rewrite rules P.

Chomsky (N Chomsky, 1956) classified the grammars into 4 levels of hierarchy : regular, context-free, context-sensitive and computably enumerable languages. These levels correspond to the amount of computational power necessary to decide whether a string belongs to the language or not. It also corresponds to constraints on the rewrite (or production) rules.

1.3.1.1.1 Regular grammars

At the bottom of this hierarchy, we find the regular grammars, which can be computed by a finite state automaton (FSA), and expressed using regular expressions. A finite state automaton is a system that can only occupy a certain number of states, and moves from one state to the other only when a triggering event occurs. FSA can be used both to generate the language by exploring the authorized transitions, and to verify whether a string of character belongs to the language by moving from one state to the next in function of the symbol contained in the string to be tested.

In the example above, there are 6 possible states. From state S0, transitions are only possible towards state S1 if a T occurs and S3 if a V occurs. Any other character causes the automaton to stop and produces a string that is not part of the language. For a grammar to be computable using a FSA, all the production rules have to be of one of the following form:

- $B \rightarrow a$ - where B is a non-terminal in N and a is a terminal symbol in Σ
- $B \rightarrow aC$ - where B and C are in N and a is in Σ
- $B \rightarrow \varepsilon$ - where B is in N and ε denotes the empty string, i.e. the string of length 0.

Regular grammars are considered to be a form of markov process (Fitch & Friederici, 2012), because the production/decision criterion only depends on the current state. However, one has to be careful in drawing the conclusions from this comparison. Similarly to what have been presented in the previous section on the predictive coding examples, Markovian transitions are computed between hidden states, which are not observable in the final string. Because a given symbol can trigger different transitions depending on the current state, and because parallel branches keep the influence of older divergences, long term dependencies may exist in the final string. For example, the inflection of the verb at the 3rd person depending on number in English can easily be recognized by a FSA, although it constitutes a long distance dependency with a variable number of elements between the subject and the verb.

1.3.1.1.2 Subregular grammars

As regular grammars can produce strings that are already complex, linguists have created sub categories of regular grammars called "subregular" grammars (Jäger & Rogers, 2012; Rogers & Pullum, 2011). A subregular language is a set of strings that can be described without employing the full power of FSAs, that is to say they can be classified by mechanisms that are simpler than FSAs, often much simpler.

1.3.1.1.2.1 Strictly local languages

The subregular hierarchy comprises at its bottom strictly local languages. Stringsets that belong to strictly local languages can be distinguished simply on the basis of which symbols occur adjacently. The length k of the sequences of symbols that have to be considered to test the belonging of a stringset to the language defines a set of k-grams that are possible, and define a strictly k-local description SL_k. The automaton that processes Local Languages are called scanners. They scan the stringset with a sliding window of length k and check in a look-up table whether the current k-gram belongs to the k-local description of the grammar. Contrary to the FSA, these automatons do not have hidden states. Grammars $(AB)^n$ that are often used in comparative psychology are SL_2: being sensitive to the 2-grams {AB, BA, ⋊A and B⋉}, where ⋊ and ⋉ note the beginning and the end of the string, is enough to distinguish strings that belong to that language from strings that do not.

The grammars defined by Chomsky do not contain a notion of probability. He used grammar to define what is possible from what isn't. In neuroimaging protocol, a standard paradigm will probe the learning of a rule by introducing violations. It implies that subjects do not consider as "grammatical" every sequence of event that is presented, but learn to expect the most likely sequences. It is possible to extend the definition of grammars by adding a probability of occurrence for each rewrite rule. In this context, the oddball paradigm could be considered a SL_1 language. Can MMN be elicited by violation of more complex k-local languages? It appears so. A MMN can be elicited by violation of the SL_2 grammar $(AB)^n$ when repetitions are introduced in the alternate sequence (ABABA*A*) (J Horváth, Czigler, Sussman, & Winkler, 2001). If the cognitive process generating the MMN was comparing bigrams in a sliding window to a lookup table, the "correct" continuation of the sequence after the violation should be ABA*A*B (AB is a frequent bigram) and the continuation ABA*A*A should be considered a violation (AA is a rare bigram). The data differ from this prediction as both sequence continuation elicit a MMN. This result suggests that either a different statistic is used, or that a larger window size (at least three items here) is taken into account. Another interesting paradigm

is the one proposed by Sussman and collaborators (Elyse Sussman et al., 1998). The SL_5 grammar $(AAAAB)^n$ was presented to subjects, without violations. No MMN was elicited by the B tone when a short stimulus onset asynchrony (SOA) was used, indicating that the SL_5 language can be learned under some circumstances. However, in the same paradigm but with a longer SOA, the rare B tone elicited a MMN. This paradigm shows that the mechanism generating the MMN is sensitive to the rate of presentation so that a potential "sliding window" is not only sensitive to the number of event but also to the time window considered.

1.3.1.1.2.2 Locally testable languages

Because they do not have internal states or memory, strictly local languages typically cannot contain any rule that requires a constraint to occur *at least* once. For example, the language that contains all strings of A and Bs that contain at least one B is not SL_k. Note that *exclusion* rules (e.g. "strings that do *not* contain any B) can be tested using SL_k descriptions.

Strictly local languages are a subcategory of a more powerful subregular type of language known as locally k-testable languages. In this type of grammar, a n-gram can be considered as a condition, that is satisfied if it occurs in a string. In addition to the set of possible k-grams, a k-testable language contains also a set of logical expressions over these conditions called k-expressions that define which combinations of the k-grams are legal in the language. A scanner for a locally k-testable language contains in addition to the look-up table, a record of the k-grams it encountered in the string. When the end of the sting is reached, the record is fed to a Boolean network which tests the k-expressions and outputs whether the string belongs to the language. Note that the testing of the k-expressions has to come after the end of the strings. Indeed, the rule "has to contain at least a B" would return a negative answer even in legal strings if the B does not occur in first position. There is no evidence that MMN is sensitive to this type of grammar.

The main limitation of locally testable languages compared to regular grammars is that they cannot contain rules that depend on the number of occurrence of each k-gram or on the relative position of each of them. Locally testable grammars are only sensitive to occurrence.

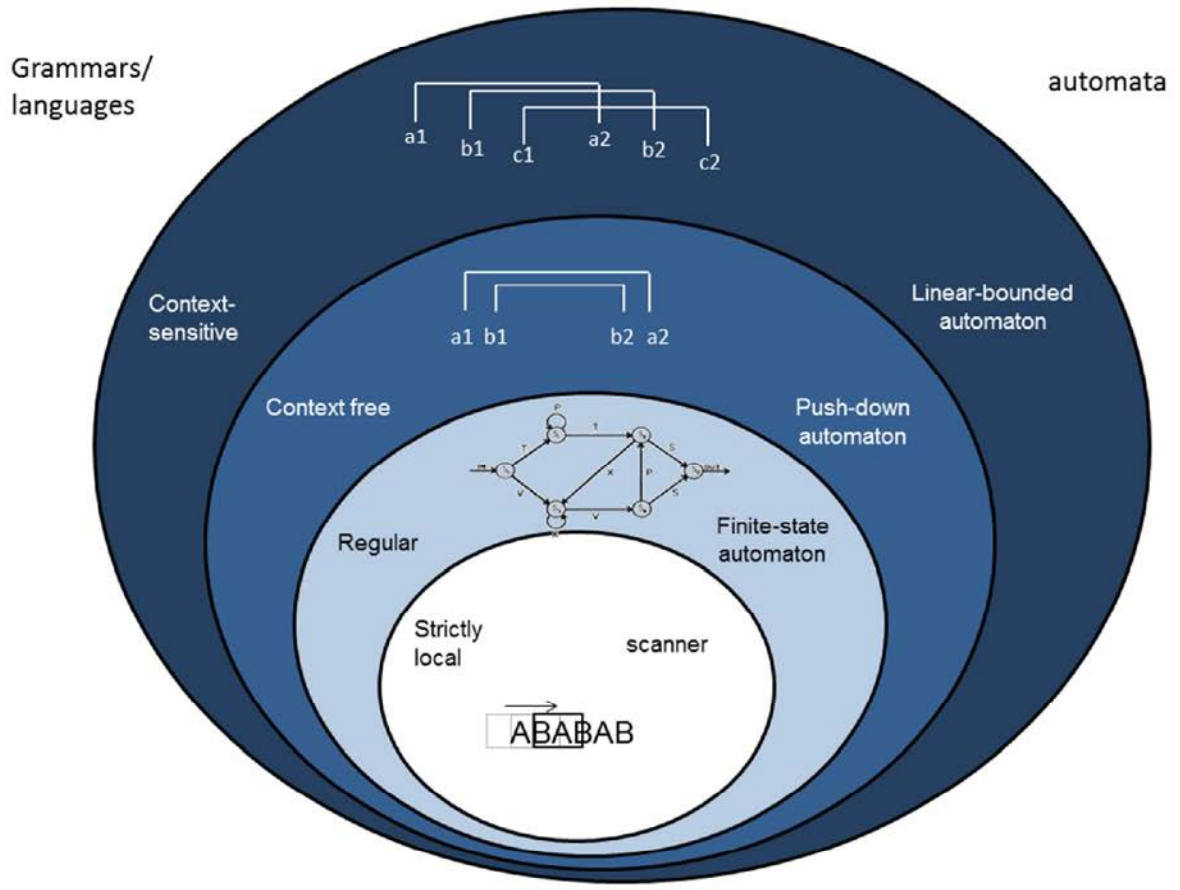


FIGURE 1.3-1: THE HIERARCHY OF GRAMMARS IN FORMAL LANGUAGE THEORY

In formal language theory, grammars and languages are organized in a hierarchical manner, so that languages at the bottom of the hierarchy are subcases of the languages that can be generated by the grammatical rules that can be used in the more complex grammars. To each grammar correspond an automaton that can be used to compute the possible strings of the language. Strictly local grammars contain only local rules that can be checked with a scanner that only considers the content of a sliding window. Regular grammars possess the full power of finite state automaton that produce the possible strings of the language by moving between hidden states in a Markovian manner. Context free languages can contain embedded structures that can be computed using a push down automaton. Context sensitive languages can be computed by linear bounded automaton and are necessary to compute cross-embeddings.

1.3.1.1.3 Context-free grammars

Context-free grammars contain regular grammars. In addition, context-free languages can also contain rules which involve embedded structures. All the production rules of context-free grammars must be of the form $A \rightarrow \gamma$, were A is a single non terminal symbol and γ is a string of terminal and/or non-terminal symbols. A typical example of a context free grammar that is not regular is the parenthesis system in mathematical expressions, which can be produced using the rules. $\{ S \rightarrow SS; S \rightarrow (S); S \rightarrow () \}$. It is impossible to imagine a FSA that can compute for an arbitrary number of parentheses the correct opening and closing, because it is necessary to keep

in memory the number of parenthesis that were open. Context free grammars are also very useful in describing phrases structure of natural languages. This type of grammar can be computed by a system with one push down stack, i.e. one memory slot that can store a potentially infinite number of items, but can only look at the last one stored. In the case of parentheses, each opened parenthesis is stacked. The parenthesis is unstacked when it is closed. A single stack is enough because there is only one class of symbol that has to be stored – the opening parenthesis- and they are closed in the inverse order they are opened.

1.3.1.1.4 Context sensitive grammars

Context sensitive grammars contain context free grammars, but accepts rules of the form $\alpha A\beta \rightarrow \alpha\gamma\beta$, where A is a single non terminal of N and α, β and γ can be strings of terminal and/or non-terminal symbols. Context-sensitive grammars correspond to linearly bounded automata. These are essentially Turing machines, i.e. FSAs with a memory tape that can perform arbitrary operations (writing and erasing symbols on the tape and moving the tape in either direction) during state transitions. The length of the available tape is not infinite, though, but bounded by a number that is a linear function of the length of the input string. The typical example of this type of grammar is a language consisting in the set of string of the form $A^nB^nC^n$.

KEY POINTS

- FLT offers a systematic hierarchical description of serial sequences complexity, based on the properties of the rules used to generate them
- The mismatch negativity has been elicited by paradigms that can be described using the simplest form of sub-regular grammar: strictly local grammars.
- FLT only considers the sequential order of symbols
- The MMN is sensitive to physical timing of events

1.3.1.2 Human natural languages and the Chomsky hierarchy

Where should we situate the human cognitive capacities in the Chomsky hierarchy? The position of human languages in this hierarchy has been established in the mid-1980's (Jäger & Rogers, 2012). Chomsky (Noam Chomsky, 1957) had already established that English was not a regular language based on the argument that it contains potentially infinite embedded structures. For example, *"the rat died"*, can contain an embedded relative in "*the rat* that the cat ate *died*", which itself could accept a relative "*The rat* the cat *the dog chased* ate *died*", and so on. But the question of whether languages where context sensitive remained open until three linguists Riny Huybregts (Huybregts, 1984), Stuart Shieber (Shieber, 1985) and Christopher Culy (Culy, 1985)

concluded independently around the same time (Shieber and Culy even pubished their results in the same issue of *Linguisitcs and Phylosophy*) that human grammars could be context sensitive. They observed that in Swiss German the dependencies between verbs and their objects are unbounded in length. However, they are not nested, but rather interleaved so that they cross each other.

Although Chomsky's hierarchy is the most systematic hierarchical description of sequential regularities, some researchers have questioned its relevance for the understanding of the underlying cognitive processes that support language processing (Perfors, Tenenbaum, & Regier, 2011; Petersson & Hagoort, 2012; Pullum & Scholz, 2010).

First, the claim that natural languages are model free relies on the idea that a potentially infinite number of levels can be embedded, which requires a form of recursive processing. However, the maximal degree of center-embedding in written language is three; and in spoken language, multiple center-embedding is practically absent (Karlsson, 2007). Native speakers tend to make errors in grammaticality judgment (Hakes, Evans, & Brannon, 1976) when the number of center embedded levels exceeds three. Even though the grammar might *potentially* accept infinite embedding, this suggests that the cognitive process used by human subjects is not capable to display the full power of a push down automaton, at least in terms of memory capacity. As a result, experimental approaches to test the hypothesis that recursion is uniquely human (Hauser, Chomsky, & Fitch, 2002) have failed to distinguish hierarchical processing from recursive processing (Abe & Watanabe, 2011; Bahlmann, Schubotz, & Friederici, 2008; Bahlmann, Schubotz, Mueller, Koester, & Friederici, 2009; Bloomfield, Gentner, & Margoliash, 2011; Gentner, Fenn, Margoliash, & Nusbaum, 2006). Moreover, this memory limitation creates the apparent paradox that humans are able to process the context sensitive structures present in some languages while not being able to process some of the center embedded structures that are considered less complex in the hierarchy.

Second, classes of the Chomsky hierarchy provide a measure of the complexity of patterns based on the structure of the algorithms (grammars, automata) that can distinguish them. However, in most case, there are multiple ways to represent the language so that stringsets can be classified correctly. For example, center embedded structures can be represented by finite state automatons as far as they don't go to an infinite number embedded levels; which seems to be the case in natural sentences. When dealing with an unknown mechanism, such as a cognitive mechanism of an experimental subject, we have no reason to think that subjects use algorithms that rely on the grammars and automata that were used to build the stimuli in the analyses they

employ in making their judgments. We know only that they can or cannot make these judgments about strings correctly.

KEY POINTS

- Context-sensitive structures exist in human languages
- A language can often be described by different grammars
- Correct classification of a sequence of symbols generated by a set of rule is not a proof that the judgment was made using the same set of rules

1.3.1.3 Using and learning Chomsky hierarchy

To understand these limitations, we have to go back to the aim of Chomsky's theory. It was mainly built to answer the question "what constitute the knowledge of language?". It argues that knowledge of a language consists in the mastery of abstract rules that allow people to distinguish between grammatical and ungrammatical sentences. It abstracted itself from other characteristics of cognition that constrain the production and comprehension of language like memory capacity limits but also lexical statistics that make some grammatical phrases more *likely* than others. That this information should be excluded was the point of Chomsky's famous sentence "Colorless green ideas sleep furiously" and the accompanying observation that, "I think that we are forced to conclude that [...] probabilistic models give no particular insight into some of the basic problems of syntactic structure" (Noam Chomsky, 1957, p17). He built a system that contained grammars that are generative and abstract in their structure. Learning such an abstract grammar based on exemplars that are not explicitly tagged as grammatical or not, but can contain ill-formed sentences is a challenge without postulating strong constraints on the underlying structure. Based on this argument of the poverty of the stimulus, Chomsky postulated that mechanisms for language acquisition were largely innate and learning consisted only in setting a few parameters of this constrained structure. However, connectionist approaches to language learning and comprehension have shown that probabilistic aspects of language bring crucial information. These networks do not try to solve the same problem as Chomsky's grammars: they try to infer meaning rather than categorize sentences as grammatical or not. In that context, probabilistic and grammatical information concur to resolve ambiguities that can occur both in the syntactic structure and lexical interpretation of a sentence (Seidenberg, 1997). For example, in the sentence "the plane left for the East Coast.", the word plane can refer to an airplane, a geometric element or a tool; and the word left could be an adjective or the past tense of the verb

leave. Deciding the meaning of the sentence requires integrating lexical and syntactic probabilistic information.

1.3.1.4 Multiple levels of complexity: an evidence for a modular processing of temporal regularity?

Advances in machine learning approaches to language acquisition and use, are associated with progress in connectionist neural networks with multiple hidden layers which capture different levels of feature regularities (DiCarlo & Cox, 2007; Hinton, 2007). Although Formal Language Theory considers the whole set of characters in a sentence, without distinguishing any intermediate parsing, it seems that language perception is itself organized at different levels (Heinz & Idsardi, 2011): we distinguish the organization of sounds into a word (the phonology) from the organization of roots and affixes into words (morphology), and the organization of words into phrases and sentences (syntax). Are all of these levels as complex? Interestingly, while the arrangement of words into sentences can be context free or even context sensitive, patterns of sounds into a word obey regular, or even subregular rules. Rules of phonology contain local dependencies such as exclusion rules regarding successive sounds. For example, the sequence of phonemes 'gling' is legal in English while the sequence 'gding' is not. In some natural languages, these exclusion rules can be at long distance (Rose & Walker, 2003). In Samala for example, a language of an Indian population from California, words cannot contain both "s" and "sh" sounds (Applegate, 2007). As a result, the word "shtoyonowonowash" is possible, but there is not word like "shtoyonowonowas". However, morphology does not rely on context-free rules.

Some authors have argued that if morphology and syntax were both relying on similar neuronal substrates for learning, there would be no reason to observe different levels in complexity for these two domains (Heinz & Idsardi, 2011). However, a long tradition in philosophy renewed by machine learning (Tenenbaum, Kemp, Griffiths, & Goodman, 2011) argue in favor of multiple learning modules specialized in difference types of regularities. Specifically, they argue that the inference problem can only be solved if learners (humans or machines) are restricted in the space of hypothesis they consider. In this view, different modules in the brain would be considering different spaces of hypotheses to explain external stimuli. The appropriate type of structure to describe a particular domain can then be discovered using a hierarchical Bayesian model that determines which part of the hypothesis space best explains the sensory data (Kemp & Tenenbaum, 2008). Bayesian approached in machine learning for the processing of speech have shown that different types of underlying structure were appropriate to learn language parsing or infer appropriate syntax (Chater & Manning, 2006). The idea that

specific brain regions could support learning of a specific type of structure is supported in the visual domain by data showing that congenitally fully blind adults that learned to interpret "soundscapes" (2D auditory transcription of a visual stimulus), ended up representing letters using the same brain area that sighted people: the visual word form area (VWFA), suggesting that the computations made by this region were more associated to the structure of perceptual data than to modality or position in the sensory hierarchy (Striem-Amit, Cohen, Dehaene, & Amedi, 2012).

KEY POINTS

- The processing of natural language does not aim at determining grammaticality. It tries to infer the correct underlying structure to determine the meaning of its elements by integrating syntactic and semantic information.
- Different aspects of natural language have different levels of complexity which is consistent with a hierarchical
- Machine learning approaches suggest that a combination of modules exploring limited hypothesis space that compete to explain the sensory inputs constitutes a powerful and efficient inference mechanism.

1.3.2 Experimental approaches to conscious and unconscious regularity processing

Can we find a dissociation between conscious and unconscious processing along the Chomsky hierarchy? Before continuing, it is necessary to clarify the definition of conscious processing. It is sometimes used to say that a representation of the regularity being processed is accessible consciously. In this context, the conscious processing of the regularity can be tested by asking for a verbal report about the nature of the regularity that was discovered or by asking subjects to bet on the accuracy of their classification (Persaud, McLeod, & Cowey, 2007). The second meaning of conscious processing implies that stimuli have to access working memory and central executive resources for the regularity to be learned. It does not imply that the regularity will be reportable by the subject, but behavioral or electrophysiological signatures of violation detection should exist when conscious access is possible. The experimental test distinguishing such processes from unconscious processes is the sensitivity of the learning performance to the manipulation of cognitive load by concurrent tasks, or attentional manipulations.

1.3.2.1 Temporal regularity learning and awareness of the rule

What kind of regularities can be discovered so that correct judgment about the conformity of a sequence to the rule can be made or appropriate behavior can be adopted, in the

absence of explicit knowledge about the nature of the rule, or even awareness that something was learned?

Learning phonology, and in particular word segmentation is one of the challenges that children have to meet, to successfully acquire language. Saffran et al. (Saffran, Aslin, & Newport, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997) showed that subjects as young as 8 month-old were indeed able to segment speech-like auditory stimuli based on transitions probability. They incidentally exposed babies, children and adults to a continuous stream of syllables for a few minutes (2 minutes for the babies, 21 minutes for the children and adults), without any prosodic markings (e.g. *bupadapatubitutibudutabapidabu*…). Syllables were organized into three-syllabic words, so that the only cues to word boundaries were the transitional probabilities between syllables pairs, which were higher within word (from 0.31 to 1.0) than between words (from 0.1 to 0.2). Even when distracted by a coloring task, the children and adults were able to classify words from non-words with better than chance accuracy. Babies were also able to distinguish the two categories even if test non-words and test words were matched for frequency in the habituation but differed in transitional probabilities (Aslin, Saffran, & Newport, 1998), ruling out the possibility that subjects were only tracking the occurrence of chunks. Interestingly, subjective reports about awareness of learning did not predict the performances of the subjects (Saffran et al., 1997). Word segmentation could therefore result from transition probability tracking between syllables. Consistently with this idea, the MMN amplitude depends on the probability of the deviant in oddball paradigms (Sato et al., 2000): the more frequent the deviant is, the smaller the amplitude of the MMN. Therefore, if MMN reflect a predictive process, this process emits probability-weighted predictions.

In Saffran's experiments, the learning occurred without any explicit instructions or intention from the subjects to learn, by mere exposure. Moreover, it occurred in people that claimed to be unaware that they learned anything and there was no effect of awareness on performance. This type of learning has been called "implicit learning". The term was coined by Reber (Reber, 1967). In a typical study, subjects first memorize grammatical strings of letters generated by a finite-state grammar. Then, they are informed of the existence of the complex set of rules that constrains letter order (but not what they are), and are asked to classify grammatical and non-grammatical strings. Subjects can classify substantially above chance, indicating that they have learned some of the underlying grammatical structure. However, they are unable to report which rule allowed them to classify the strings. Reber described this results as a "peculiar

combination of highly efficient behavior with complex stimuli and almost complete lack of verbalizable knowledge about them" (p. 859).

Most artificial grammar learning (AGL) research focused on regular grammars and found typically better than chance classification. A few studies used context free grammars (Rohrmeier, Fu, & Dienes, 2012) and showed also better than chance performance without observing conscious access to the set of rules that was used. These results suggest that the complexity of sequences according to the Chomsky hierarchy does not determine conscious access to the rules.

1.3.2.1.1 Consciousness of the rule and symbolic representations

One of the main limitations of the AGL protocol comes from the difficulty to establish what type of criterion was used by subjects to perform the grammaticality judgment. In particular, models that store the occurrence of chunks of sequences met in the training exemplar can perform as well as Reber's subjects. Experimental manipulation of the frequency of bigrams in test and training exemplars showed that this parameter does affect the grammaticality judgment of subjects (Kinder & Assmann, 2000). To prove that subject could learn the underlying structure of the sequence and not just the frequency of bigrams, researchers have attempted to show that subjects can transfer the structure of the grammar to new stimuli and use it to classify strings that do not contain any of the bigrams presented in the training examples. Transfer of knowledge to a new set of stimuli in AGL paradigms have yielded mixed results. Significant transfer is typically observed but with a drop in performance compared to within same set generalization. The interpretation of these results remained inconclusive (Dienes & Altmann, 1997).

One of the hypotheses to explain the incomplete transfer performance was that consciousness of parts of the rule was acquired during learning – typically repetition patterns – and that only that part was transferred when the stimuli changed. Indeed, a more contrasted result is observed in attempts to test transfer of repetition patterns to new sequences: only subjects that are to report awareness of the regularity achieve successful transfer in serial reaction time tasks (P. F. Dominey, Lelekov, Ventre-Dominey, & Jeannerod, 1998). This argues in favor of the idea that the type of structure that is learned implicitly in AGL learning in SRT is mainly statistical or at least to concern surface properties of the stimulus while conscious access allows more flexibility by processing more abstract objects.

The abstractness of a representation can be classified according to three levels (Buchler, 1955; Deacon, 1997; Nieder, 2009): iconic representations are truthful to at least some of sensory features of the object they represent. Indexical representations allow arbitrary mapping

from one domain to another, so that each correspondence has to be specified. Symbolic representations do not only represent mappings between different domains. They also rely on the *relations* between the objects of one domain so that the mapping from one domain to the other can be *inferred* based on the similarity of the relations for new objects.

Symbolic processing based on repetition pattern can already be found in babies. Marcus (Marcus, Vijayan, Bandi Rao, & Vishton, 1999) familiarized seven-month-old infants with a continuous stream of syllables without prosody. The syllables were organized into three-syllables "words" that were formed using a very simple grammar relying on "algebraic" rules: in the ABB condition, the babies listened to a stream of words were the two last syllables were identical and different from the first one (*gatititalalagagigilinana*…). In the test phase, they were presented with stimuli constituted from new syllables that were never presented during the habituation. These new stimuli could either continue to follow the previous grammar (CDD) or follow a different grammar (CDC). The infants were able to discriminate the two types of test stimuli. Given that no correspondence between the syllables was instructed, the transfer of the grammar from one set of syllables to the other can only be possible if it was inferred on the bases of the relations of repetition among the stimuli.

1.3.2.1.2 Implicit conditioning

The reportability of the rules that are learned was also investigated in classical conditioning.

The acquisition of classical delay eyeblink was shown to be behaviorally similar whether subjects were aware or not of the relation between the two stimuli (R. E. Clark, 1998; Robert E. Clark, Manns, & Squire, 2002; Manns, Clark, & Squire, 2002). However, in a minimal variant of the paradigm, the CS is made shorter, so that there is some interval between the end of the CS and the air puff while the time between the onset of the CS and the US remains unchanged. This paradigm is called trace conditioning, because it requires that a sensory trace of the CS be maintained during the temporal gap between the two stimuli for learning to occur. In trace eyeblink conditioning, significant conditioning behavior could not be observed in subjects that were not able to report verbally the association between the CS and the US.

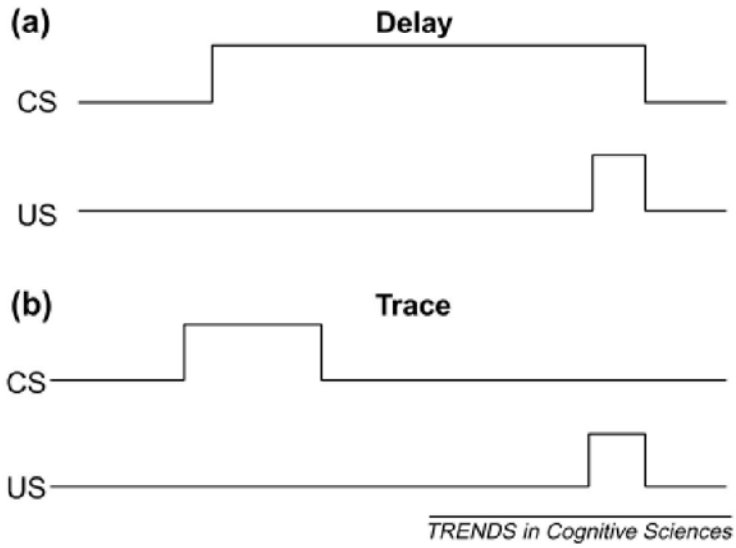


FIGURE 1.3-2: TRACE AND DELAY CONDITIONING.

(a). In delay conditioning the US occurs during or at the end of the CS. (b). In trace conditioning, a delay is introduced between the end of the CS and the US. The delay between the onset of the CS and the US is not necessarily longer in trace conditioning than in delay conditioning. From (Robert E. Clark et al., 2002)

KEY POINTS

- Complexity according to the Chomsky hierarchy does not predict whether rules that are learned can be accessed or not.
- Access to symbolic representations can allow transfer across stimuli sets.
- Transfer of symbolic rule is not observed when subjects are not able to report the rule.
- In eyeblink conditioning, when a delay is added between the CS and the US, the efficacy of the conditioning and the awareness of the relation are tightly correlated

1.3.2.2 Temporal regularity learning and central resources

The other important axis of research regarding the interplay between structure learning and conscious access looks at which type of learning is affected by manipulation of attention or task interference. In other words, learning that requires conscious access and conscious processing *of the stimuli*. Here the criterion for learning is not meta-knowledge but behavioral manifestations of learning and their resilience to attentional manipulations.

In the previous paradigm of eyeblink conditioning (R. E. Clark, 1998), both awareness and behavioral acquisition of the association was affected by a distractor task in trace conditioning. On the contrary delay conditioning was not affected by the distractor task. Thus, it

seems that bridging the temporal gap requires attentional resources, while in the absence of a delay the conditioning is more automatic. The automaticity of delay eyeblink conditioning is supported by data in rabbits showing that decerebrate animals (i.e. after removal of cerebrak cortex, basal ganglia, limbic system, thalamus and hypothalamus) but with intact brainstem and cerebellum, exhibited retention of normal delay conditioning (M D Mauk & Thompson, 1987). In humans, delay conditioning is impaired in patients with cerebellar lesions (Daum et al., 1993; Topka, Valls-Solé, Massaquoi, & Hallett, 1993) or brainstem lesions (Solomon, Stowe, & Pendlbeury, 1989) but intact in amnesic patients with damages that include the hippocampus. On the contrary, trace eyeblink conditioning is affected not only by lesions to the cerebellum (D S Woodruff-Pak, Lavond, & Thompson, 1985) and its afferences, but also by damages to the hippocampus and neocortex. Acquisition and retention of trace conditioning was severely disrupted by hippocampal damages in rabbits (Kim, Clark, & Thompson, 1995; Moyer, Deyo, & Disterhoft, 1990) and rats (Weiss, Bouwmeester, Power, & Disterhoft, 1999). It was also affected by damages to the prefrontal cortex, including anterior cingulate cortex (ACC), in rabbits (Kronforst-Collins & Disterhoft, 1998; Powell, Skaggs, Churchwell, & McLaughlin, 2001; A. P. Weible, McEchron, & Disterhoft, 2000). Consistently with these data, trace conditioning is affected in human patients with damages to the hippocampus. Interestingly, this impairment is proportional to the trace interval: while conditioning was mildly affected for short delay traces (500ms) it was more strongly affected for longer traces (1000ms) (McGlinchey-Berroth, Carrillo, Gabrieli, Brawn, & Disterhoft, 1997). In this type of patients delay conditioning with CS durations of 750ms is unaffected, which implies that bridging the temporal gap is the crucial challenge in trace conditioning. Moreover, the hippocampus seems to be implicated only transiently in the acquisition of trace conditioning (Kim et al., 1995) while damages to prefrontal cortex affect performances at any stage of the process(McLaughlin, Skaggs, Churchwell, & Powell, 2002). Consistently with these data, trace fear conditioning was found to be affected by a working memory distractor task (Carter, Hofstotter, Tsuchiya, & Koch, 2003 but see Carrillo, Gabrieli, & Schaaf, 2000). At the neuronal level, prefrontal cortex activity increases during the trace interval and trace conditioning is affected by disruption of delay activity by injection of GABAA agonist, NMDA antagonists and delay specific optogenetic silencing in medial prefrontal cortex of rats (Gilmartin & Helmstetter, 2010; Gilmartin, Miyawaki, Helmstetter, & Diba, 2013), showing causally the implication of working memory delay activity in bridging the temporal gap. Overall, data suggest that cerebellum and PFC cooperate to respectively determine the precise timing and bridge the temporal gap between he CS and the US (Kalmbach, Ohyama, & Mauk, 2010; J. J. Siegel, Kalmbach, Chitwood, & Mauk, 2012). The ACC has been implicated

in attentional processes that signal the relevance of the CS when it is paired with the US and associated to context specific maintenance of CS during the delay (a P. Weible, Weiss, & Disterhoft, 2007).

Consistently with this results in conditioning, data from serial reaction time tasks point also to a crucial role of attention and working memory for the learning of long distance dependencies. In SRT, long distance dependencies between elements distant of up to 6 positions have been argued to be acquired (Remillard & Clark, 2001; Remillard, 2008, 2010), given enough trials (up to several tens of thousands for 6^{th} order dependencies!). Long distance dependencies learning was also observed in babies (Gómez, 2002): 18 month-old infants were able to discriminate 2 languages constituted of three-syllables words presenting identical first-order transition probabilities but different non-adjacent dependency between the first and last syllables of the words(e.g. aXb, cXd and eXf belong to language1 while aXe , bXf , and cXd belong to language2). However, Curran and Keele (1993) showed that acquiring second- or higher order transitions is blocked—in terms of both acquisition and use—by performing a demanding secondary task, whereas learning first-order transitions is unaffected, suggesting again that maintenance of information would be key for long distance learning to occur, and for that learning to be exploitable.

KEY POINT

- The most reliable dissociation between processes that are affected and processes that are resilient to attentional manipulation and cognitive load in the need for maintenance of information across a temporal gap

1.3.3 Simultaneous but opposite expectations for conscious and unconscious processing

Are conscious and unconscious levels of processing tightly coupled? Is low-level unconscious processing modulated by higher-level conscious expectations? Can we detect consciously a deviance without low level response to a violation of regularity?

Although the MMN and the P300 are often correlated, in particular in the oddball paradigm (K. C. Squires et al., 1976), they can be dissociated in paradigms where conscious expectations are manipulated to differ from low level regularities. In a bimodal protocol, if a rare visual stimulus signals the impending occurrence of a rare tone, whereas a frequent visual stimulus signals that a frequent tone will be presented, the rare but predictable auditory stimuli do not elicit a P300 response while the MMN response is maintained (Ritter, Sussman, Deacon, Cowan, & Vaughan, 1999a). This dissociation translates at the behavioral level by the suppression

of the behavioral effects of distraction by rare tones (E Sussman, Winkler, & Schröger, 2003). Other manipulation of conscious expectations have yielded the same results (János Horváth, Winkler, & Bendixen, 2008; Rinne, Antila, & Winkler, 2001).

On the other side, the local-global paradigm exposed in the first section of this introduction (Bekinschtein et al., 2009) showed that a P300 response and conscious report of violation detection could be elicited in the absence of a MMN response; and reproduced the absence of P300 response in the presence of a predicable MMN.

These results argue in favor of a local computation of temporal expectations at both levels.

1.4 Neuronal properties for the implementation of a temporal predictive process

The previous sections have led us to consider that predictions and predictions errors seem to be a general mechanism for the learning and processing of temporal regularities, either consciously or unconsciously. We have identified temporal gaps as a recurring characteristic of regularities that cannot be learned or exploited without access to working memory.

In this section I will examine some of the properties of the neuronal bases that could support learning of temporal regularity. I will first focus on the dynamics of learning in MMN paradigms and relate them to the main neuronal bases of learning. Then I will shortly give an overview of the key properties of the representations that can be found in auditory cortex. Finally I will review the type of neuronal codes that could support prediction and learning of timing and/or identity of future stimuli.

1.4.1 Temporal dynamics of learning

How fast are new regularities learned?

Most data about the dynamics of regularity learning for the MMN come from oddball paradigms. Using frozen oddball sequences with multiple reversal between standard and deviants, the adaptation of the response of individual neurons, and of the ERP to the standard stimulus follows an exponential dynamic, with time constant of a few tens of seconds for the whole population and up to a few seconds or stimuli for single neurons (Costa-Faidella, Grimm, Slabu, Díaz-Santaella, & Escera, 2011; Ulanovsky, Las, Farkas, & Nelken, 2004). In roving paradigms(N. Cowan, Winkler, Teder, & Näätänen, 1993), a tone is repeated for a variable number of times,

until the presentation of a different tone that becomes then the new standard tone, until the next change. The first tone of each new sequence represents a deviant for the previous one, and a MMN can be observed by computing the difference between the response to the deviant and the response to the last standard tone. The MMN was already present after series of standard of length 2 and was maximal after only 6 repetitions of the new standard (Haenschel, Vernon, Dwivedi, Gruzelier, & Baldeweg, 2005), which represent time constant even shorter than in the oddball paradigm. One of the key differences between the oddball with periodic inversion of the standard and deviant and the roving paradigm is that in the first paradigm, the transition probabilities are inverted from one inversion to the other, while in the second, the otherwise certain transition probability given the previous tone is only violated once at the end of a period, then when this tone is presented again after some time, the previous transition was never violated again.

The importance of local history was illustrated by Squires (K. C. Squires et al., 1976) showing that the amplitude of the MMN-P3 depends strongly on the past few items. The result was reproduced at the single neuron level in A1(Ulanovsky et al., 2004). The influence of local patterns in random sequences translates into behavior (Huettel, Mack, & McCarthy, 2002; Schvaneveldt & Chase, 1969). Interestingly reaction times are not only sensitive to repetition patterns, but also to alternate patterns. The sensitivity to alternate patterns emerges however more slowly than sensitivity to repetition. MMN presents also this property: repetition of a stimulus in an alternate sequence (ABABABA*A*..) elicits a MMN (J Horváth et al., 2001).

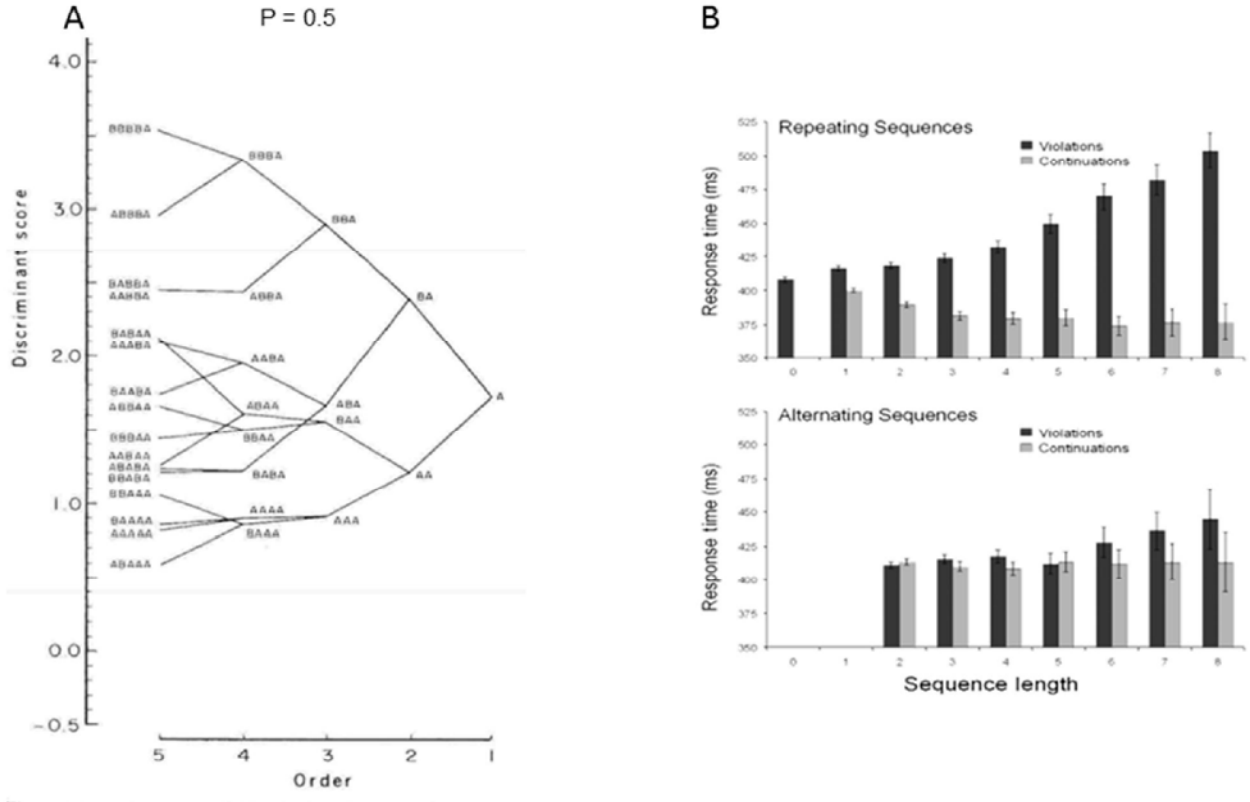


FIGURE 1.4-1 EFFECT OF RECENT HISTORY ON TEMPORAL REGULARITY PROCESSING

(left) In an oddball paradigm, effect of the recent history on the amplitude of the MMN-P300 response. Probability of the two tones A and B are .5/.5. The amplitude of the response to a tone A is shown in function of the identity of the last 1, 2, 3 4 of 5 stimuli. It is larger when previous stimuli where mostly Bs than when most previous stimuli where As. Adapted from Squires (1976). (right) Behavioral effect continuation and violation of runs of consecutive repetitions or alternations in a serial reaction time task. Subjects were instructed that targets appeared at random in one of two locations. Runs of consecutive repetitions or alternation could occur randomly. The difference between the reaction time to continuation or violation of the regularity increased with the number of repetition of the pattern. From Huettel (2002)

KEY POINTS

- MMN appears within a few presentations of a repetitive stimulus
- The oddball paradigm confounds potentially multiple types of long term and short term learning mechanisms
- The alternation of two stimuli is also learned within a few trials

1.4.2 Neuronal mechanisms of sequence learning

1.4.2.1 Short term plasticity

Short term synaptic plasticity (H Markram, Wang, & Tsodyks, 1998) reflects neurotransmitter vesicle dynamics at the synapse. When an action potential reaches the

presynaptic terminal a proportion of the vesicles present in the terminal and containing neurotransmitters are released, depleting the stockpile. As a result, if another spike reaches the terminal before the full stock of vesicle is recovered, the proportion of vesicle released represent a smaller quantity of neurotransmitters and generates a smaller post synaptic potential. The time constant of stock recovery and the proportion of vesicle used for each potential determine the dynamics of short term depression. The time constant vary from tens of milliseconds to seconds (Varela et al., 1997), which is consistent with the dynamics of learning in the oddball paradigm. Note that the oddball paradigm does not allow disambiguation between mechanisms that are linked to short term plasticity of "stimulus specific adaptation" (SSA); i.e. mechanisms that will only be able to adapt to the repetition of the feature coded by the neuron, and mechanisms that would rely on prediction of the identity of the next stimulus, whether it is identical to the previous one or follows a more complicated pattern. However, short term plasticity does not explain the emergence of an MMN to repetition in an alternate sequence, unless very specific neuronal encoding schemes are postulated. Crucially, any "learning" effect associated to short term plasticity should attenuate rapidly, and eventually vanish in the absence of stimulation within a few tens of seconds.

1.4.2.2 Long term plasticity

Long term plasticity refers to more durable changes in synapse efficacy. The most famous rule of long term plasticity is Hebb's rule (Hebb, 1949):

"Let us assume that the persistence or repetition of a reverberatory activity (or "trace") tends to induce lasting cellular changes that add to its stability.… When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased."

This rule was proposed to explain associative memory, as a form of supervised learning, where a "teacher" signal (equivalent to the US in conditioning) would lead a neuron B (corresponding to CS) to fire whenever a nearby neuron A (corresponding to a conditioned response) fires. After teaching, applying Hebb's rule, neuron A would be able to trigger activity in neuron B on its own. This idea is consistent with data from incidental learning showing a strong similarity between the associations learned under reinforcement conditions and the associations learned by mere exposure.

Ahissar (Ahissar et al., 1992) designed a protocol to specifically assess associative training-induced changes in the correlation strengths of functionally coupled neuronal pairs separated by several hundred microns across the cortex. One neuron of a pair functioned as a conditioned stimulus (CS) neuron, and the second functioned as a conditioned response (CR) neuron. The unconditioned stimulus (US) was an auditory stimulus capable of driving the CR neuron and guiding the monkey's performance on an auditory task. The activity in the CS neuron triggered the US and therefore activity in the CR neuron. Cross correlations before and after training revealed an increase in the coupling between the CS and CR neuron. This change in efficacy lasted for a minute after the end of the associative pairing. These results revealed apparently powerful plasticity changes in excitatory intracortical interneuronal connections obeying a Hebb rule (Buonomano & Merzenich, 1998), but were dependent on the concurrent performance of a rewarded task. A similar cell conditioning protocol paired the preferred stimulus of a cell with silencing of its activity and the less preferred stimulus with high activity by injecting respectively inhibitory and excitatory currents when the stimuli were presented (Frégnac, Shulz, Thorpe, & Bienenstock, 1992). After this "conditioning" the receptive field has shifted durably towards the conditioned receptive fields, even in the absence of reward signal.

At the cellular level, synaptic plasticity in cortex and hippocampus has been described to be tightly dependent on the timing of spikes in the presynaptic and postsynaptic neurons (Levy & Steward, 1983). This spike timing dependent plasticity (STDP) was shown to be dependent on the back propagation the post-synaptic spike (Henry Markram, Lübke, Frotscher, & Sakmann, 1997) and the specific critical window for STDP was described in detail by Bi and Poo(G.-Q. Bi & Poo, 1998). Consistently with the associative learning, STDP typically shows a depression if the post synaptic spike occurs before the presynaptic spike; and a potentiation of the synapse if the post synaptic spike follows the presynaptic spike. As a result, the synapse will only be reinforced if the presynaptic activity "predicts" the post synaptic activity. At the behavioral level, the same is true: backward conditioning where the CS follows the US does not result in an association.

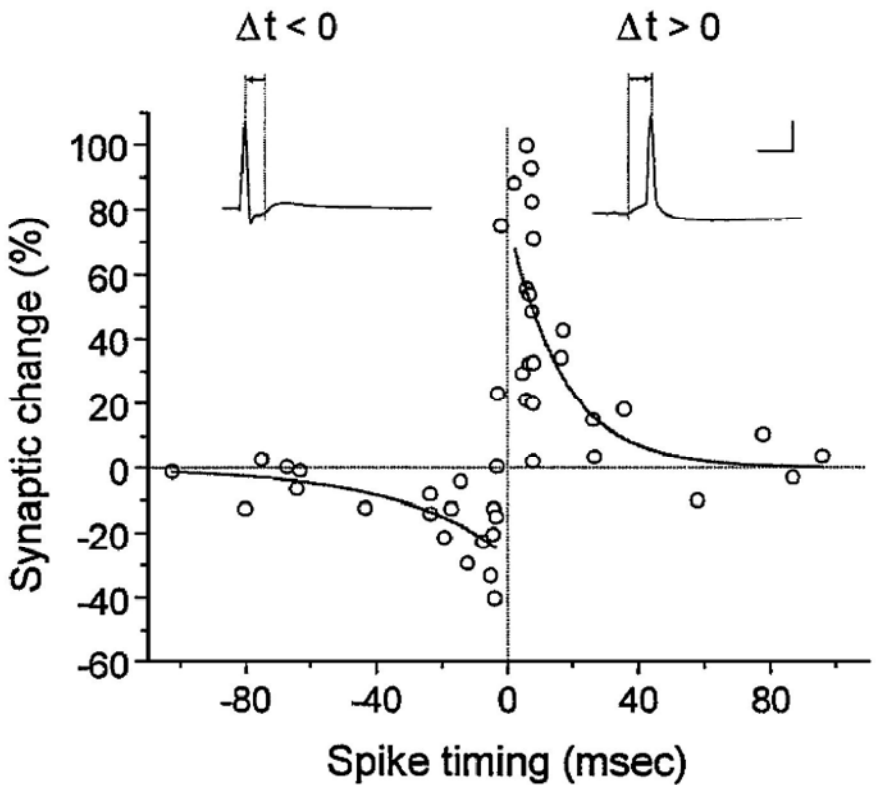


FIGURE 1.4-2: SPIKE TIMING DEPENDENT PLASTICITY

Percentage of change in the amplitude of post synaptic currents 20 min after a repetitive correlated spiking between a presynaptic neuron and a post synaptic neuron with a delay Δt (and exponential fits). Maximum increase in synaptic efficacy is observed when the presynaptic spike occurs with 40ms before the postsynaptic spike. From (G. Bi & Poo, 2001; G.-Q. Bi & Poo, 1998)

R. P. Rao & Sejnowski (2001) showed how a STDP learning rule can allow a network of neocortical neurons to predict an input a few milliseconds before the input's expected arrival. This mechanism was also proposed for associative learning across temporal distance using delay lines to transform time into a spatial coding (G. Bi & Poo, 2001).

Moreover, STDP relies on the back propagation of the post synaptic spike in the dendritic tree and is NMDA-dependent (G.-Q. Bi & Poo, 1998). Interestingly, the MMN is abolished by the injection NMDA-receptor antagonists (Javitt et al., 1996), suggesting that NMDA-dependent plasticity could be crucial in MMN generation.

KEY POINTS

- Short term synaptic habituation presents dynamics consistent with learning dynamics observed in the oddball paradigm
- Long term Hebbian plasticity can be induced by associative pairing of a presynaptic and post synaptic activity driven by a third input.
- The long term plasticity depends on the relative timing between pre and post synaptic spikes
- Long term plasticity is NMDA-dependent.

1.4.3 Neuronal codes in auditory cortex

The primary auditory cortex is located in the temporal lobe, in the Heschl gyrus. It receives thalamocortical projections from the thalamic medial geniculate body. The principal functional organization of primary auditory cortex is its tonotopic arrangement. The receptive field of each neuron can be determined using pure tones with varying frequency and loudness. The characteristic frequency is defined as the frequency for which the neuron responds at the lowest sound level. The characteristic frequencies are organized in an orderly manner with neurons at the anterior part of A1 being sensitive to high frequencies and at the posterior part to low frequencies in cats (Reale & Imig, 1980), rats (Rothschild, Nelken, & Mizrahi, 2010) and humans(Moerel, De Martino, & Formisano, 2012).

However, the tonotopy is not the only dimension mapped onto the auditory cortex. Topographic organization of sharpness of the frequency response area (Schreiner & Mendelson, 1990), threshold and response latency have been observed. Tonotopy only exists in mice at a relatively gross scale, while the local organization is much more complex (Rothschild et al., 2010). However, this organization remains patchy as reflected by a high noise correlation at the local scale. This complex organization with highly correlated noise is consistent with the existence of partly overlapping networks encoding topographically different dimensions of the input.

Temporal coding of a sound is on average characterized by a strong response to the onset and offset of the stimulus with a quieter activity in the intermediate period (Qin, Chimoto, Sakai, Wang, & Sato, 2007). Cells can respond either to only the onset or the offset in a particular frequency band, or to both, potentially with different preferred frequencies. Given this encoding scheme, the onset and offset of a stimulus can be considered as separate events that need to be

predicted. Consistently with this idea, MMN can be elicited by changes in duration or in interstimulus intervals (Ford & Hillyard, 1981; Näätänen et al., 1989).

1.4.4 Predicting timing and identity in neuronal networks

What neuronal networks support the memory for timing or for the identity of past stimuli?

1.4.4.1.1 Prediction of timing

A predominant hypothesis in the psychological literature on conditioning has been that timing could rely on a centralized internal clock (Treisman, 1963) in which an oscillator beats at a fixed frequency generating tics that can be detected by a counter. Cerebellar patients are known to present deficits in timed interval comparison and production (Ivry & Keele, 1989) and both trace and delay conditioning are strongly affected by cerebellar lesions(Diana S Woodruff-Pak & Disterhoft, 2008). The cerebellum has therefore been proposed to fulfill the role of this general purpose timing mechanism. However, evidence for a role of the cerebellum in timing comes mainly from lesion studies and no precise data exist (Michael D Mauk & Buonomano, 2004). The contingent negative variation (CNV) (Walter et al., 1964) has been considered as a neurophysiological support of the internal clock model. This evoked potential is linked to temporal expectancy and develops in the interval between two stimuli, when the first predicts the second. The CNV takes the form of a ramping activity. The development of the CNV requires a learning phase of about 30 repetitions of the paired stimuli and requires the execution of a task – which can be purely mental – in response to the second stimulus. When the expected duration is longer than the expected one, the ramping activity peaks at the expected interval and drops (Macar & Vidal, 2003, 2004). CNV amplitude also correlates with variability in interval production (Macar, Vidal, & Casini, 1999). However, the CNV is dependent on attention and it is unlikely that the same mechanism is used for pre-attentive stimulus timing prediction where no ramping activity can be found.

1.4.4.1.2 Dynamic attractors: tracking time *and* identity of past stimuli

Another encoding scheme comes from investigation of associative learning at the neuronal level. It is widely believed that the precise timing of the CR in classical conditioning could be achieved by relying on a memory trace that transforms temporal information into a spatial code that can be exploited by associative plasticity mechanisms(G. Bi & Poo, 2001). Bi and Poo found in hippocampal neuron cultures that repetitive paired-pulse stimulation of a single neuron for brief periods induces persistent strengthening or weakening of specific polysynaptic

pathways in a manner that depends on the interpulse interval. These changes can be accounted for by correlated pre- and postsynaptic excitation at distant synaptic sites, resulting from different transmission delays along separate pathways. This mechanism is apparent to a "delay line" where information about time is coded in the pattern of activity in the network. Synfire chains are typical examples of such networks: units are arranges in population connected in a feedforward manner, so that activity propagates through the network, establishing a linear relationship between the neuronal population active at one given time and the delay since the input was feeded to the first population of the network. Such networks have been shown to have the capacity to support learning of complex spatiotemporal patterns (Grossberg, 1970). Support for such a encoding scheme was also described in auditory cortex slices where stimulation of one neuron triggers a reproducible sequential activation of neurons arranged in a non-topographical manner (Buonomano, 2003). Therefore, temporal information could be multiplexed into auditory cortex without interfering in a systematic way with the other topographical maps.

A reproducible sequence of activation is not a proof of a feedforward organization, and the cortical connectivity is thought to be mainly recurrent. Learning from networks that produce reproducible activity trajectories across time while presenting a more realistic encoding scheme has been the subject of intensive investigation (Buonomano & Laje, 2010; Michael D Mauk & Buonomano, 2004). Recurrent networks have attracted interest under different names in cognitive neuroscience (*temporal recurrent networks*, Dominey, Arbib, & Joseph, 1995), in computational neuroscience (*liquid computing*, Maass, Natschläger, & Markram, 2002) and in machine learning (*echo state networks* Jaeger, 2001). The general approach consists in having a recurrent network of interconnected units that respond to an input by a reverberating activity that follows a rich but reproducible trajectory. Readout units are then trained to use this activity to produce a desired output. These networks can learn very complex input/output mapping including delayed responses because the network dynamic carries information about identity and timing of past inputs, although in a less linear manner than delay lines. A general weakness of this approach is however its tendency to be chaotic and sensitive to noise: two slightly different inputs will tend to produce very different spatiotemporal dynamics. Moreover, the dynamics elicited by one given input might vary dramatically if realistic synaptic noise is introduced. However, recent advances to generate locally stable trajectories could be used to successfully simulate robust timing and motor patterns (Laje & Buonomano, 2013).

1.4.4.1.3 Working memory: selective maintenance of identity…and timing ?

1.4.4.1.3.1 Capacity of working memory: slots or variable precision?

A second weakness of recurrent networks is the sensitivity to interference from other inputs that effectively reduces the temporal window that is reliable enough to be usable for learning. Working memory is thought to overcome this limitation by selecting the information that is maintained, thus reducing the interference problem. The dominant model of working memory proposes that a fixed number of discrete items can be held in discrete "slots" that can each contain arbitrary complex information about one object or chunk (Buschman, Siegel, Roy, & Miller, 2011; Fukuda, Awh, & Vogel, 2010; Luck & Vogel, 1997). The exact number of slots was originally estimated to about seven (G. Miller, 1956), but more recent observations led to a current estimate of four (Nelson Cowan, 2001, 2010; Luck & Vogel, 2013; Edward K Vogel & Machizawa, 2004), although it has been argued to be even more reduced (Olsson & Poom, 2005). One important limit of behavioral measures of working memory capacity is the difficulty to control possible encoding strategies of the subjects that reduce the number of chunks effectively encoded compared to the experimenter belief (Bor, Duncan, Wiseman, & Owen, 2003). However, the number of item encoded in working memory in spatial memory task correlates with the amplitude of an ERP called the contralateral delay activity (CDA). Interestingly, the amplitude of the CDA saturates for a number of items that correspond to the psychophysics estimate of memory capacity (Edward K Vogel & Machizawa, 2004). Moreover, this potential is specifically proportional to the number of items that are recalled, and not to the distractors that have to be ignored (Edward K Vogel, Mccollough, & Machizawa, 2005). Crucially, in subjects presenting a "low memory capacity", the CDA was sensitive to distractors. The filtering efficiency, which measures how sensitive the CDA is to distractors, is strongly correlated memory capacity, suggesting that limitation in the memory capacity is not so much a limitation in encoding capacity that a deficit in selective maintenance of the appropriate information. A neurophysiological mechanism for working memory capacity limits has been proposed by computational models that utilize neural oscillations as the primary representational scheme for information being held in working memory. Indeed, the power of oscillatory activity in the theta range (4-10Hz) has been showed to increase specifically during retention delay in working memory tasks (Raghavachari et al., 2001) and to correlate with the number of items being successfully retained (Jensen & Tesche, 2002; Sederberg, Kahana, Howard, Donner, & Madsen, 2003). A decoding approach suggests that reactivation of memory content following a theta rhythm (Fuentemilla, Penny, Cashdollar, Bunzeck, & Düzel, 2010). Theta rhythm also modulates the amplitude of gamma oscillation (20-35 Hz) during working memory tasks (M. Siegel, Warden,

& Miller, 2009) and prefrontal cortex neurons tend to fire at specific phases of theta and gamma oscillations that differ from one object to the next if multiple items are maintained (H. Lee, Simpson, & Logothetis, 2005). Integrating these results, neurobiological models(Lisman & Idiart, 1995; Raffone & Wolters, 2001) have proposed that neuronal assemblies fire synchronously at specific phases of the theta cycle accomplishing both binding and segregation of features into objects in working memory. Limitations in spiking correlation resolution would explain the limitation to three or four items.

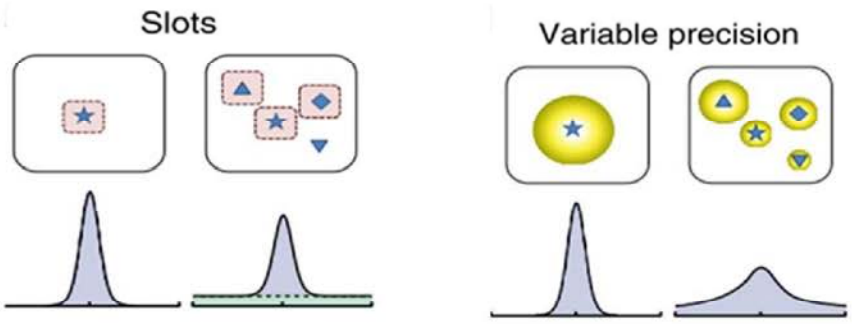


FIGURE 1.4-3: SCHEMATIC REPRESENTATION OF THE MAIN MODELS OF WORKING MEMORY

(left) (a) In the slot (or item limit) model of working memory, each visual item is stored in one of a fixed number of independent memory slots (here, 3) with high resolution (illustrated, by narrow distribution of errors around the true feature value of a tested item). When there are more items than slots, one or more items are not stored and the slot model predicts that errors in report of a randomly chosen item will be composed of a mixture of high-precision responses (right, blue component of distribution corresponds to trials when the chosen item received a slot) and random guesses (green component corresponds to trials where it did not get a slot). (b) Variable precision: working memory precision varies, from trial to trial and item to item, around a mean that decreases with increasing number of items as a result of limited resources. This model predicts that recall errors will be made up of an infinite mixture of distributions (assumed normal) of different widths. Variability in precision could stem from variability in resource or from bottom-up factors From (Ma, Husain, & Bays, 2014)

Recently, an alternative model gained increasing interest which proposes that memory does not encode information into discrete slots, but distributes attentional resources between a potentially infinite number of items, regulating the precision of their encoding, i.e. the resistance of the representation to noise across time (Ma et al., 2014; van den Berg, Shin, Chou, George, & Ma, 2012). This model accounts best for the behavior of memory precision in tasks that test the precision of memory for analogical properties of the stimulus that may vary along a continuum like color, length or orientation of stimulus. Memory capacity has been shown to be limited to one object in a task that requires precise representation a non-categorical stimulus, while it was consistent with previous estimates of four items if the shapes could be categorized (Olsson & Poom, 2005).

The most prominent correlate of working memory is increased sustained activity in lateral prefrontal cortex during maintenance (Fuster, 1971; Goldman-rakic, 1995). However, in recent years, multivariate pattern analysis showed that information about the maintained stimuli could be decoded from sensory cortices, even though it did not involve increased activity during delay, but rather persistence of a specific pattern of activation (Harrison & Tong, 2009; Linden, Oosterhof, Klein, & Downing, 2012; Offen, Schluppeck, & Heeger, 2009; Pasternak & Greenlee, 2005; Postle, Druzgal, & D'Esposito, 2003; Riggall & Postle, 2012). Although fMRI decoding study show typically poor decoding of content in prefrontal cortex, neuronal recordings in that region show that neurons do encode identity specific information about the object which suggest that this result points to lack of sensitivity of the method to the encoding patterns in prefrontal cortex rather than to a lack of specificity of representations. However, in addition to displaying coarse selectivity for the sensory features of the object being maintained (Rainer, Rao, & Miller, 1999), lateral prefrontal cortex also exhibits selectivity for task relevant categorization (D J Freedman, Riesenhuber, Poggio, & Miller, 2001; David J Freedman, Riesenhuber, Poggio, & Miller, 2002; E. K. Miller, Freedman, Wallis, Trans, & Lond, 2002; E. K. Miller, Nieder, Freedman, & Wallis, 2003; Roy, Riesenhuber, Poggio, & Miller, 2010). Recent data show that depending on the task that followed the maintenance delay. Specifically, they observed stronger decoding during the maintenance of visual than categorical properties in posterior fusiform cortex, whereas the opposite was true in lateral prefrontal cortex (S.-H. Lee, Kravitz, & Baker, 2013). Taken together with the data that led respectively to the proposal of the slot model and the variable-precision model of working memory, these results suggest that two dissociated systems can be used to maintain information: a slot-based system that relies mainly on prefrontal cortex for categorical information, and a more precision-dependent system that relies on sensory cortex for analogical sensory information. The choice between these two systems is goal-dependent.

1.4.4.1.3.2 Resistance to noise and distractors across time: the role of dopamine

How are stimuli selected to be encoded in working memory and how is memory content protected from distractors? Several lines of evidence implicate dopamine in the selective maintenance in working memory. First, the degradation of dopaminergic neurons in patients suffering from Parkinson disease and in animal models impaired by lesions show deficits in working memory (Lange et al., 1992; Miyoshi et al., 2002; a. M. Owen et al., 1992). Low doses of dopamine agonists can enhance working (Servan-Schreiber, Carter, & Bruno, 1998). Dopamine agonist or antagonist injected directly in prefrontal cortex affect specifically working memory while leaving other functions intact (Sawaguchi & Goldman-Rakic, 1991, 1994).

Neuroimaging studies linked the activity in basal ganglia to working memory and in particular to efficiency of distractor filtering (D'Esposito, Postle, Ballard, & Lease, 1999; Lewis, Dove, Robbins, Barker, & Owen, 2004; McNab & Klingberg, 2008; Menon, Anagnoson, Glover, & Pfefferbaum, 2000). Computational models (M. J. Frank, Loughry, & O'Reilly, 2001; Gruber, Dayan, Gutkin, & Solla, 2006a; O'Reilly & Frank, 2006) have proposed that modulation of dopamine in the basal ganglia allows gating of working memory and improves the reliability of representations across time. These models suggest that dopamine's involvement in affective and reward processing endows this gating with specificity to motivational salience, allowing for example the learning of a relevant task-dependent strategy of gating to complete complex memory tasks.

1.4.4.1.3.3 Time in working memory

How is time represented in working memory? One of the most commonly used class of models of working memory tend to rely on a bistable activity of a whole population recurrently connected that maintain a stable code across time, thus losing the temporal information about the time of occurrence of the maintained stimulus(Camperi & Wang, 1998; Wang, 1999, 2001). However, neuronal data tend to show more complex dynamics with possibly multiple representations of time. First, in delayed match to sample (DMS) task, individual neurons in PFC showed step-like increment in their firing rate for each non-match item presented (E. K. Miller, Erickson, & Desimone, 1996). This coding scheme gives an event-based information about time with a linear relation between the firing rate and the number of stimuli met since the encoding of the stimulus in working memory. Second, a sophisticated analysis of somatosensory delay activity in prefrontal cortex led author to conclude that the representations of timing and stimulus identity are concurrently maintained by separate mechanisms, while sharing a common anatomical substrate, creating orthogonal representations of the two features in a high dimensional space (Machens, Romo, & Brody, 2010). Third, an analysis of the dynamics of the temporal code in visual delayed match to sample showed a rapid transformation of the code in the first few hundreds of millisecond towards an increasingly stable code (Stokes et al., 2013). This result is compatible with a logarithmic precision in the representation of time as would be expected in a functionally feedforward network with slow synaptic dynamics (Goldman, 2009).

1.4.4.1.4 A mechanism for prediction without precise timing

Miller found that in delayed paired associate task activity in prefrontal cortex was progressively evolving from a retrospective activity encoding the feature of the sample object, to a prospective activity sensitive to the feature of the expected target (Rainer et al., 1999). This

evolution did not seem to be time specific, but could be the result of Hebbian association between the two stimuli of a pair leading to the transition from one code to the other (Mongillo, Amit, & Brunel, 2003). How is prediction compared to input to determine the match or non-match response? Prefrontal neurons have been described to show match-enhancement of their response when the incoming input corresponds to working memory content in match to sample task or in delayed paired associate task (E. K. Miller et al., 1996; Rainer et al., 1999). This mechanism can be used to test the conformity of stimulus identity to expectations coded in PFC without having to predict the timing of the stimulus in addition to its identity (Engel & Wang, 2011).

KEY POINTS

- Time and identity of past stimuli can be encoded dynamically in neuronal activity.
- Working memory maintains the identity of a limited number of items
- Dopaminergic systems are involved in the selective maintenance of information in working memory
- Neuronal mechanisms in prefrontal cortex could allow prediction of the identity of the next stimulus without its timing

CHAPTER2

INTRODUCTION TO THE PERSONAL CONTRIBUTIONS

The review presented in the previous chapter showed that understanding the interplay between conscious and unconscious processing of temporal sequences is an exciting fundamental question that has a long history in cognitive science. It is also a clinical challenge as understanding the specificity of conscious processes could result in more sensitive and more specific diagnostic tools for non-communicative patients. Neuronal and behavioral responses associated with violations of temporal regularities and their differential sensitivity to attentional and awareness manipulations represent a promising direction for further investigations. Neuronal responses to violation of regularities, in particular, represent a privileged way to probe regularity processing in patients that cannot behave or communicate. Responses to violation of temporal regularities are widely believed to be correlates of prediction errors that reveal predictive processes in the brain. We will adopt this framework in the following chapters. By predictive coding we will refer to the general framework that proposes that the brain build a generative model of the input and use it to generate predictions about the incoming stimuli, without restricting ourselves to the efficient coding principle.

I chose to focus on two main axes: *i)* understanding the organization of unconscious processing of temporal regularities and the neuronal constraints that limit its computational capacities, and *ii)* examine under which conditions properties of conscious processing can overcome these limits.

The review of the roles of predictions and prediction errors in behavior and neuronal data highlighted two main purpose for prediction errors in addition of carrying the information about unpredicted stimuli: a behavioral role driving attentional resources towards new unexpected events and a role in learning as a "teacher" for acquisition of a predictive model. The fact that MMN occurrence is neither necessary nor sufficient to orient attention towards novel events rules out the behavioral contribution of MMN. Moreover, the fact that MMN is not abolished by top-down expectations argues in favor of a local computation of states dynamics, as

predicted by the free-energy framework. I will test the hypothesis that prediction error can be used in local cortical circuit as a guide for the learning of accurate temporal predictions.

- Can we build biologically plausible predictive coding model that learns to predict future stimuli based on past exposure to a temporal regularity and reproduce the main properties of MMN?
- Can this model give better predict new data than the fresh-afferent model?

In the **Chapter 3**, we will propose a biologically plausible model of MMN based on the predictive coding framework. In particular, we will test the hypothesis that the main properties of MMN can be reproduced by a predictive model that uses prediction error as a "teacher" signal. We will then use the prediction of this model to propose an experimental protocol that can disentangle the predictive coding from the other main account of MMN: the "fresh afferent hypothesis".

Moreover, the large range of feature that elicit MMN responses engage us to think that MMN reveals a computational principle of the cortical hierarchy rather than a A1-specific process. However no evidence exists to date showing simultaneous predictions and predictions errors at multiple hierarchical levels. The hierarchy of regularities proposed by (Bekinschtein et al., 2009) constitute an ideal paradigm to test this hypothesis.

- Is the computation that generates MMN reproduced at multiple hierarchical levels?

In **Chapter 4** we will adapt the hierarchical sequence used by (Bekinschtein et al., 2009) to test the hypothesis that the computational module described in chapter3 is duplicated in multiple hierarchically organized levels.

Finally conscious access and access to working memory are thought to be two very similar phenomenons. Interestingly, the maintenance of information over temporal gaps proved to be an important parameter that distinguishes the temporal regularity which require an access to working memory to be learned.

- What are the computational benefits and constraints associated with processing of stimuli in working memory?

The **Chapter 5** will focus on one property of conscious processing: the capacity to hold past stimuli in working memory for an arbitrary long time. We will examine the computational challenges and the processing abilities associated with this particular form of memory.

CHAPTER3

ARTICLE 1: A NEURONAL MODEL OF PREDICTIVE CODING ACCOUNTING FOR THE MISMATCH NEGATIVITY

Published in Journal of Neuroscience (2012)

3.1 Introduction to the article

3.1.1 Goal of the article

In the introduction chapter we reviewed the three main models proposed to account for the mismatch negativity (MMN): *i)* a memory based model characterized by a separate comparison module that generate a mismatch response if the stimulus differs from a memory trace that stored the repetitive aspects of the previous stimuli (Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001), *ii)* an habituation model that relies on short term depression of repeatedly stimulated synapse implemented by May & Tiitinen (2009) and *iii)* a predictive coding interpretation which states that the MMN represents a prediction error signal (Friston, 2005). Only the habituation model has been implemented at the neuronal level.

In this chapter, I propose a biologically plausible implementation of a predictive model that learns to predict future stimuli based on past temporal regularities. I compare the simulated responses of the model to different types of sequences, to the known properties of the MMN. Moreover, May & Tiitinen (2009) argued that all MMN properties could be accounted for by synaptic habituation – given neuronal populations with the appropriate selectivity – making the predictive coding interpretation of the MMN unnecessarily complex. I used the prediction of the model to build an experimental protocol in which predictions from the predictive model and from the habituation model are qualitatively different.

3.1.2 Choice of the methods

I chose to implement the model of the mismatch response at the spiking neuron level so that realistic synaptic STDP rules could be used for learning. I chose to use Izhikevich equations

(Izhikevich, 2003) to model spiking neurons, as they are one of the best model available to reproduce the behavior of different types of neurons. However, the results presented in this paper are not crucially dependent on the choice of spiking neuron model and could be reproduce using exponential integrate and fire models implemented in standard spiking neural network simulators such as pyNEST or Brian.

The model is used to develop an experimental protocol in which the habituation model and our predictive coding model make qualitatively different predictions. This protocol is then tested using magnetoencephalography (MEG). While EEG records the electric field MEG records the magnetic field at the surface of the scalp. Because the magnetic field interacts little with biological tissues, the MEG presents the advantage over EEG that it is spatially more precise, while being as good temporally. Magnetic fields propagate in a direction orthogonal to electric fields. The topography and polarity of the mismatch response is therefore different in MEG compared to EEG. The magnetic equivalent of the MMN (MMNm) consists in a larger response to deviant than to standard stimuli over the temporal electrodes between 100 and 200ms after the onset of the sound.

3.1.3 Reference

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

3.2 Article

3.2.1 Abstract

The mismatch negativity (MMN) is thought to index the activation of specialized neural networks for active prediction and deviance detection in auditory cortex. However, a detailed neuronal model of the neurobiological mechanisms underlying the MMN is still lacking, and its computational foundations remain debated. We propose here a detailed neuronal model of auditory cortex, based on predictive coding, that accounts for the critical features of MMN. The model is entirely composed of spiking excitatory and inhibitory neurons interconnected in a layered cortical architecture with distinct input, predictive and prediction error units. A spike-timing dependent learning rule, relying upon NMDA-receptor synaptic transmission, allows the network to adjust its internal predictions and use a memory of the recent past inputs to anticipate on future stimuli based on transition statistics. We demonstrate that this simple architecture can account for the major empirical properties of the MMN. These include a frequency-dependent response to rare deviants, a response to unexpected repeats in alternating sequences (ABABAA…), a lack of consideration of the global sequence context, a response to sound omission, and a sensitivity of the MMN to NMDA receptor antagonists. Novel predictions are presented, and a new magneto-encephalography experiment in healthy human subjects is presented that validates our key hypothesis: the MMN results from active cortical prediction rather than passive synaptic habituation.

3.2.2 Introduction

Since it was first described at the end of 1970's , the Mismatch Negativity (MMN) has been largely used in theoretical and clinical research (for review see Näätänen, 2003). It was first recorded by EEG in the context of the oddball paradigm. In the most frequently used version of this paradigm participants are instructed to listen to repeated occurrences of one sound, called the standard. This monotony is disrupted at rare moments by the presentation of a different sound, called the deviant. The difference in the responses evoked by deviants and standards takes the form of a broadly negative waveform at the top of the scalp, which peaks between 100 and 200ms after the onset of the sound. MMNs can be elicited by differences in sound frequency, duration (Näätänen et al., 1989), amplitude (Näätänen et al., 1987), or inter stimulus interval (Ford & Hillyard, 1981). MMN is resistant to manipulations of attention and states of wakefulness (Sculthorpe, Ouellet, & Campbell, 2009) even though these parameters can modulate its amplitude. An analog of MMN was described in visual (Pazo-alvarez et al., 2003; Tales et al., 1999), olfactive (Krauel et al., 1999; Pause & Krauel, 2000) and somatosensory (Kekoni et al.,

1997; Shinozaki et al., 1998) modalities, supporting a broad computational significance of MMN as a shared and automatic brain mechanism responsive to stimulus novelty.

MMN is frequently interpreted in terms of predictive coding (T. S. Lee & Mumford, 2003; R. P. Rao & Ballard, 1999) assuming that the brain does not respond passively to incoming inputs, but learns the inputs regularities and uses that knowledge to actively predict what should happen next. The auditory system would acquire an internal model of regularities in auditory inputs, including abstract ones, that are used to generate weighted predictions about the incoming stimuli (Näätänen, Jacobsen, & Winkler, 2005; Paavilainen, Jaramillo, Näätänen, & Winkler, 1999; Winkler, 2007). If these predictions differ from the actual stimulus, it results in a mismatch signal.

While mathematical models of predictive coding have been proposed (Garrido, Kilner, Kiebel, & Friston, 2007; Kiebel, Daunizeau, & Friston, 2008, 2009), including some attributing distinct functions to the various cortical layers (Friston, 2005), none of them has yet led to a precise neuronal implementation of the generators of the MMN, in terms of realistic receptors, synapses and spiking neurons. Nor has there been a systematic comparison of the models' predictions with actual experimental results. Furthermore, not everyone accepts the predictive interpretation of MMN. May and Tiitinen (2009) argue that synaptic habituation (reduction of the EPSP following repetitive stimulation of the same synapse) is sufficient to explain all of the properties of the MMN and thus, that there is no need to postulate an elaborate prediction and comparison mechanism.

Here we propose a neuronal network model, devoid of synaptic habituation but comprising a detailed implementation of predictive coding, accounting for a large amount of data on the MMN. The model leads to the distinction of several processes that contribute to the observed event-related responses, and makes new predictions, one of which is tested here with magneto-encephalography (MEG).

3.2.3 Material and methods

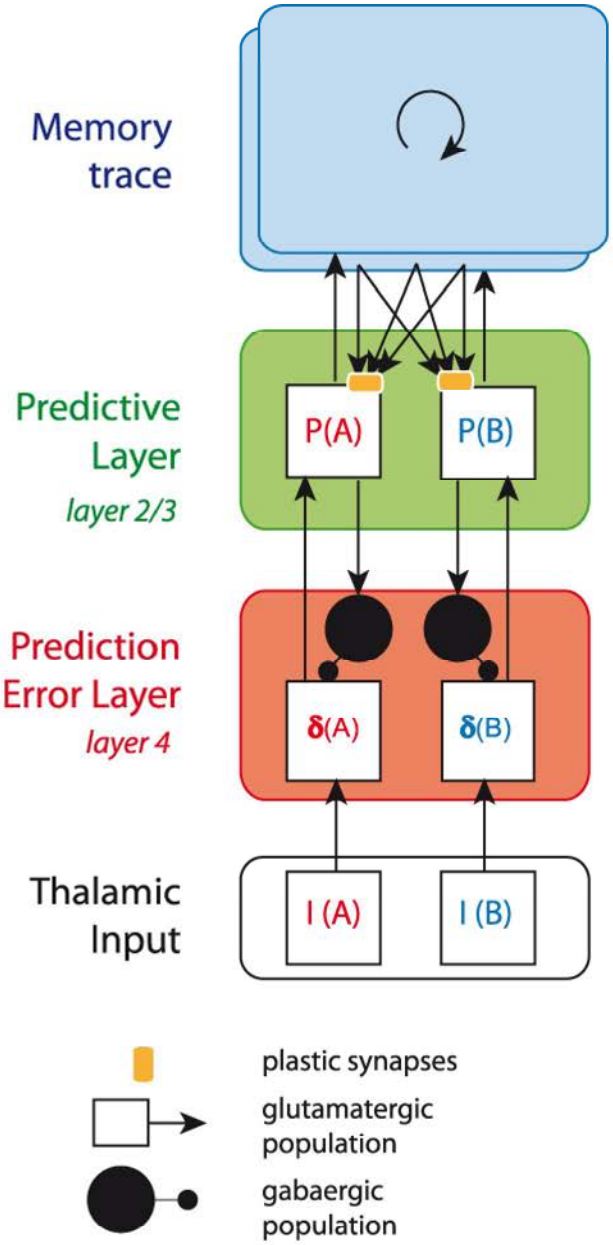


FIGURE 3.2-1: SCHEME OF THE PREDICTIVE CODING MODEL FOR TWO SOUNDS.

For each layer two subpopulations are modeled that respond respectively to the frequencies of sounds A and B. Prediction error activity in layer 4 is the result of the difference between thalamic inputs and predictive activity arising from the supragranular layer, whose sign is inverted through inhibitory interneurons (black circles). Prediction error is then fed back in order to adjust the activity of predictive populations. Dynamic predictions are made possible in the model because predictive units send and receive projections with a recurrent network serving as a short-term memory. NMDA dependent plasticity adjusts the synaptic weights onto predictive units until their dynamics matches that of the inputs and therefore miminizes the prediction error.

3.2.3.1 Network architecture

The proposed neuronal network aims at modeling the response of primary auditory cortex to incoming sounds. Figure 3.2-1 shows an implementation of the model for an input composed of two pure tones, hereafter called A and B. Each column of the network represents a cortical column with its thalamic input responding maximally to one of the two frequencies of the input. The two frequencies A and B are supposed to be different enough to activate only one of the two columns.

In each column, three populations of neurons are simulated. The essential component of the model is the population of neurons involved in prediction, which we propose to be part of the supragranular layers of the cortex. This population constantly tries to anticipate the upcoming auditory inputs. A prediction of sound A consists in an increase in the population firing rate coding for this stimulus.

At every moment, the continuously variable predictions arising from the predictive populations of neurons are compared to the incoming inputs. This comparison is achieved at the level of a population of neurons called the "prediction error" population, which receives two sets of inputs : excitatory inputs coming from the thalamus and conveying the current sensory stimulus, and inhibitory inputs that reflect the activity of the predictive population. Through this scheme, whenever the thalamic input is not cancelled by predictive signals, the prediction error population fires. The activity of the prediction error population is transmitted to the predictive population as a feedback and this error signal is used to adapt the internal model of this population (see the description of the learning rule further below). We show in the result section that this error signal may account for the MMN effect.

The predictive population needs to build an internal model of the regularities of the incoming stimulus in order to form relevant predictions. We propose that this model is based on learning the statistical temporal dependencies linking the stimuli within the past few hundred milliseconds. A memory of the recent past is needed to achieve such a goal. This memory has to keep the trace of two properties: the identity of the past inputs and the time elapsed since they occurred. We choose to model this function in the simplest manner possible, using a delay line for each frequency, where activation propagates linearly from one neuron to the next as a function of time. The relevance of this model will be discussed later.

Memory neurons are connected to both predictive subpopulations so that predictions of one frequency (A) can be based on the recent occurrence of a sound of the other frequency (B).

The internal model of the predictive population is built by adapting the synaptic weights linking the memory neurons and the predictive populations.

3.2.3.2 Detailed implementation

All subpopulations are composed of 40 neurons, except for delay lines that are composed of 400 excitatory neurons and 100 inhibitory neurons. All populations receive an external input I_{ext} that is Gaussian noise of mean equal to zero and variance equal to 2.5 for input neurons, 2 for predictive neurons and prediction error neurons, 3.8 for interneurons.

By default, mean synaptic weight between two excitatory neurons is $w_{EE} = 1.4$, between an excitatory and an inhibitory neuron $w_{EI} = 4.5$, and between an inhibitory and excitatory neuron $w_{IE} = 22$. If a presynaptic neuron is excitatory, w_{EI} *or* w_{EE} is the weight for AMPA mediated currents. An NMDA-receptor dependent current is added whose weight w_n is 20% of the AMPA synapse. The synaptic weights are drawn from a Gaussian distribution with a variance of 20% of the mean. These parameters allow a reliable transmission of activity from one population to the other in absence of other inputs, while avoiding unrealistic synchrony of neurons due to excessive homogeneity in the parameters.

The probability of a connection between thalamic inputs and prediction error populations is p = 0.9. The probability of a connection between predictive populations and interneurons and between interneurons and prediction error neurons, is p = 0.55. Synapses between predictive populations and memory neurons were initialized with weight w = 0.4 and variance of 20% with a probability of connection of 0.5. Connectivity between layers is consistent with neocortical local circuitry data (Thomson & Lamy, 2007).

3.2.3.3 Spiking neuron model

We used spiking neurons whose membrane potential is computed according to Izhikevich (Izhikevich, 2003) equations:

$$\frac{dv}{dt} = 0.04v^2 + 5v + 140 - u + I_{syn}$$

$$\frac{du}{dt} = a(bv - u)$$

Where v is the membrane potential and u a membrane recovery variable. The neurons fire if their membrane potential reaches 30mV and is then reset as follow :

$$if\ v \geq 30mV, \qquad then \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases}$$

The parameters for excitatory (resp. inhibitory) neurons were :

a = 0.02 (resp. 0.06 + 0.04*rand^2),b = 0.2 + 0.04*rand^2 (resp. 0.2), c = -65+10*rand^2 (resp. -65), d = 8 – 2*rand^2 (resp. 2), where rand is a random number drawn from a uniform distribution between 0 and 1. These parameters correspond respectively to regular spiking neurons for excitatory neurons and fast spiking ones for inhibition (Izhikevich, 2003).

AMPA, NMDA and GABA synaptic currents are modeled according to Brunel & Wang (Brunel & Wang, 2001).

$$I_{syn}(t) = I_{AMPA}(t) + I_{NMDA}(t) + I_{GABA}(t) + I_{ext}(t)$$

With

$$I_{AMPA}(t) = g_{AMPA}(V(t) - V_E) \sum_{j=1}^{C_E} w_j^{AMPA} s_j^{AMPA}(t)$$

$$I_{NMDA}(t) = \frac{g_{NMDA}(V(t) - V_E)}{\left(1 + [Mg^{2+}]\, exp(-0.062\, V(t)/3.57)\right)} \times \sum_{j=1}^{C_E} w_j^{NMDA} s_j^{NMDA}(t)$$

$$I_{GABA}(t) = g_{GABA}(V(t) - V_I) \sum_{j=1}^{C_I} w_j^{GABA} s_j^{GABA}(t)$$

Where $s_j^{receptor\ type}$ is a variable describing the opening dynamic of the receptors : AMPA and GABA receptors have instantaneous opening and close up with time constants $\tau_{AMPA} = 2ms$ and $\tau_{GABA} = 10\ ms$.

I_{ext} is an additional current that accounts for the sensory inputs from cochlear neurons which are not implemented in the model.

$$\frac{ds_j^{AMPA}(t)}{dt} = \frac{s_j^{AMPA}(t)}{\tau_{AMPA}} + \sum_k \delta(t - t_j^k)$$

$$\frac{ds_j^{GABA}(t)}{dt} = \frac{s_j^{GABA}(t)}{\tau_{GABA}} + \sum_k \delta(t - t_j^k)$$

NMDA receptors have slower dynamics with opening time constant $\tau_{NMDA,rise} = 2\ ms$ and closing time constant $\tau_{NMDA,decay} = \ 100\ ms$, $\alpha = 0.5\ ms^{-1}$.

$$\frac{ds_j^{NMDA}(t)}{dt} = \frac{s_j^{NMDA}(t)}{\tau_{NMDA,decay}} + \ \alpha x_j(t)\left(1 - s_j^{NMDA}(t)\right)$$

$$\frac{dx_j(t)}{dt} = \frac{x_j(t)}{\tau_{NMDA,rise}} + \sum_k \delta\left(t - t_j^k\right)$$

3.2.3.4 Synaptic plasticity

To internalize the statistical regularities that relate past activity to present stimuli, we implemented synaptic plasticity only between memory neurons and predictive subpopulations. We used a spike-timing dependent plasticity (STDP) rule (G. Bi & Poo, 1999) producing conditioning association :

If a post synaptic spike at time t that follows a presynaptic spike :

$$\Delta w_{pré,post} = c_p(\ I_{Ca^{2+}} - Th)\ \exp\left(\frac{t - t_{spike\ pré}}{\tau_p}\right)$$

If a pre-synaptic spike follows a post synaptic spike that occurred at time t :

$$\Delta w_{pré,post} = -c_p(\ I_{Ca^{2+}} - Th)\ \exp\left(\frac{t - t_{spike\ post}}{\tau_p}\right)$$

In addition, we used a long-term depression (LTD) rule, that induces a small depression of synapses whenever the presynaptic neuron spikes. This rule is in agreement with experimental observation that synapses tend to depress when they do not elicit postsynaptic spike (Debanne, Shulz, & Frégnac, 1998):

$$\Delta w_{pré,post} = -c_d\ \ \delta(t - t_{spike\ pré})$$

The parameters used for the simulations presented in this paper are $c_p = 60,\ \tau_p = 30ms\ c_d = 100\ et\ Th = 2.5$.

We verified that our qualitative results were largely independent of the fine tuning of the parameters. I_{Ca2+} is a calcium current mediated by NMDA receptors. This current is taken equal to I_{NMDA} for each predictive neuron.

3.2.3.5 Simulations:

For each simulation, a new network was generated following the above probabilistic connectivity rules. Each condition was simulated on 5 to 10 different networks, plotted results are averages over all simulations. Inputs were an additional I_{ext} current with amplitude 1.9, injected in the thalamic subpopulation coding for the sound corresponding to the stimulus presented. The input for each simulation was created by pseudo-randomization of a set of trial containing the desired proportions of standard and deviant stimuli. The randomization was made so that two deviants were never consecutive. Standard stimuli immediately following deviant stimuli were removed from analysis.

Various paradigms were simulated by modifying the sequence of A and B inputs in different stimulus blocks. The classical oddball paradigm was simulated as a sequence of 2000 tones, where 5%, 10%, 20% or 30% of the tones were B tones (deviants) and other tones were A, with a stimulus onset asynchrony (SOA) of 200ms. The connectivity matrix was saved after each tone, 100ms after the onset of the tone. The mean connectivity matrix that we report in Figure 3.2-4 represents the average connection strength between the memory neurons and the predictive population. It was obtained by averaging these matrices over each subpopulation of predictive neurons and over all tones except the first 200. Alternate sequences were composed of 1500 pairs of alternating tones (ABAB…, ISI=200 ms). The reproduction of the local-global paradigm (Bekinschtein et al., 2009; Wacongne et al., 2011) was made by starting with 20 standard sequences (100% AAAAB; ISI=150 ms) followed by 100 sequences comprising 70% standards (AAAAB), 20% deviants (AAAAA), and 10% omissions (AAAA). For the omission effect, a simulation of 1500 pairs of sounds (AA, ISI=200 ms) was also performed, with 10% of pairs replaced by single tones (A). We compared this to the response to 500 single tones (A).

3.2.3.6 MEG experiment

3.2.3.6.1 Participants

Five healthy volunteers (3 males, 2 females, mean age: 22) with no neurological or psychiatric problems were studied. All participants gave their written informed consent to participate to this study, which was approved by the local Ethical Committee.

3.2.3.6.2 Auditory Stimulation

Pairs of 50-ms-duration sounds were presented via headphones with an intensity of 45 dB and 200 ms stimulus-onset asynchrony (SOA) between sounds. Each sound was a pure sinusoidal tone (either 800 Hz -low, or 1600 Hz -high).

Sounds were organized in two blocks. In each block, the frequent pair, comprising two distinct sounds (AB), was first presented 10 times, with 1s SOA between pairs. 120 pairs were then presented, with SOA varying between 10 and 20 s, and with 70% of frequent AB pairs, 10% of rare AA pairs, 10% of rare BB pairs, and 10% of rare BA pairs. The identity of the A and B tones was swapped between blocks. The pairs were pseudo-randomized so that two rare pairs were never consecutive. Frequent pairs following immediately a rare pair are excluded from the analysis. All stimuli were presented using E-prime software v1.1 (Psychology Software Tools Inc.).

3.2.3.6.3 MEG/EEG recordings

Measurements were carried out with the Elekta Neuromag® MEG system (Elekta Neuromag Oy, Helsinki, Finland) installed at the NeuroSpin center (Saclay, France), which comprises 204 planar gradiometers and 102 magnetometers in a helmet-shaped array. ECG as well as EOG (horizontal and vertical) were simultaneously recorded as auxiliary channels. MEG and auxiliary channels were low-pass filtered at 330 Hz, high-pass filtered at 0.1 Hz and sampled at 1 KHz. The head position with respect to the sensor array was determined by four head position indicator coils attached to the participant's scalp. The locations of the coils and EEG electrode positions were digitized with respect to three anatomical landmarks (nasion and preauricular points) with a 3D digitizer (Polhemus Isotrak system®). Then, head position with respect to the device origin was acquired before each MEG/EEG recording session.

Each participant was recorded for 1h15: 2 sessions of about 33min duration separated by a short resting period. Participants were asked to keep their eyes open and to avoid eyes movements by staring at a fixation cross. Participants were instructed to pay attention to the auditory stimuli. Importantly, although subjects were attending to the stimuli, which may generate additional attention-dependent components such as N2b, these components typically do not contribute to MEG signals (Alho et al., 1998). At the end of the recording, a question list was submitted to the participant. This list aimed to determine which regularities the participant was able to report after recording.

3.2.3.6.4 Post-processing

Artefacts arising from outside the sensor array, such as those stemming from limb movement or other ambient magnetic disturbances, were greatly reduced by the signal space separation method (SSS) (Taulu et al., 2004). Gradiometers and magnetometers with amplitudes continuously exceeding 3000 fT/cm² and 3000 fT respectively were set as bad channels and excluded from further analysis. SSS correction, head movement compensation and bad channels correction were applied using the MaxFilter Software (Elekta Neuromag®).

A principal-component analysis (PCA) was used for PCA-based removal of EEG and EOG artifacts. Signal was averaged around artefacts for each channel type (EEG, axial and longitudinal gradiometers, and magnetometers) and a PCA was performed. Main components were saved.

The rest of the preprocessing was performed using Fieldtrip software (http://fieldtrip.fcdonders.nl/). Trials were epoched for each trial type between 200ms before and 800ms after the onset of the first sound. A low-pass filter at 40Hz was applied and PCA correction of cardiac and EOG artifacts was performed using the PCA components previously computed. The trials were base-line corrected using the first 200 ms of the epoch.

After visual rejection of jump and pronounced trend artefacts, the data were averaged per condition and per participant. The latitudinal and longitudinal gradiometers were combined by computing the mean square root of signal at each sensor position.

3.2.3.6.5 Statistics

Statistics were performed using Fieldtrip cluster-based statistics. To examine differences between experimental conditions, paired t-tests were performed with a threshold set at $p = 0.05$. Significant samples were clustered in connected sets on the basis of temporal and spatial proximity. Cluster statistics were calculated by taking the sum of t-values in every cluster. To obtain a p-value corrected for the size of the search space (time X sensors), a Monte Carlo Method was used to evaluate how extreme the cluster statistics of the two conditions were compared to random partitions of the samples. The proportion of random partitions that resulted in larger cluster statistics than the observed one was the p-value. The threshold was fixed to corrected $p = 0.05$.

Statistics on the difference between the frequent AB condition and the rare AA condition were computed between 0 and 300 ms after the onset of the second sound.

3.2.3.6.6 Response amplitude

The amplitude of the response to each of the two tones was defined as the average response over all magnetometers in the time window of the peak response for each sound (ie between 95 and 125ms after the onset of the first tone and between 135 and 160ms after the onset of the second tone). The amplitudes were normalized for each subject by the response to the first sound averaged over all conditions.

3.2.4 Results

3.2.4.1 Oddball paradigm and MMN

We first simulated the response of the network to the classical oddball paradigm. For this simulation, the network received as inputs two stimuli A and B, corresponding to sounds of frequencies distant enough to activate non overlapping populations of neurons. The input neurons were supposed to be selective only to the onset of the sound and were thus stimulated by an extra input current on input populations during 10ms. The first stimulus ("sound A") was presented most of the time (standard tone), and the other one ("sound B") more rarely, with a parametrically variable frequency (deviant tone).

The left panels of Figure 3.2-2 show the response to the standard and deviant tones, averaged over all analyzed presentations, in the specific case where the deviant has a 10% probability of occurrence. One can immediately observe that both the firing rates and the synaptic currents of the prediction and prediction-error neurons (but not the sensory neurons) are higher on deviant than on standard trials. The detailed neuronal mechanisms of this mismatch effect are the following. First, note that the predictive population coding for the sound A starts firing shortly before the occurrence of both standard and deviant sounds (top panel, red curve). This activity originates from the excitatory post synaptic currents coming from the memory neurons: the network predicts the forthcoming occurrence of a sound A. This activity inhibits the prediction error layer via an interneuron population. If a sound A is actually presented it cancels most of the excitation coming from thalamic inputs, resulting in a minimal prediction-error response. As seen in Figure 3.2-2, only a small proportion of prediction-error neurons still fire on standard trials, primarily due to stochastic fluctuations in the onset and strength of delay and predictive neurons, which therefore fails to full cancel the incoming signal. On the contrary, when a deviant sound B is presented, the prediction of an A sound does not cancel the input for a B sound. This results in a large prediction-error response which is relayed to the predictive subpopulation coding for B in order to adapt the predictive model. It forces the neurons of the predictive layer to discharge and causes a large NMDAr-dependent current that results in

NMDAr-dependent plasticity. This plasticity leads to an adaptation of the internal model of the network, reinforcing the synapses coming from the delay lines that discharged just before the prediction error signal.

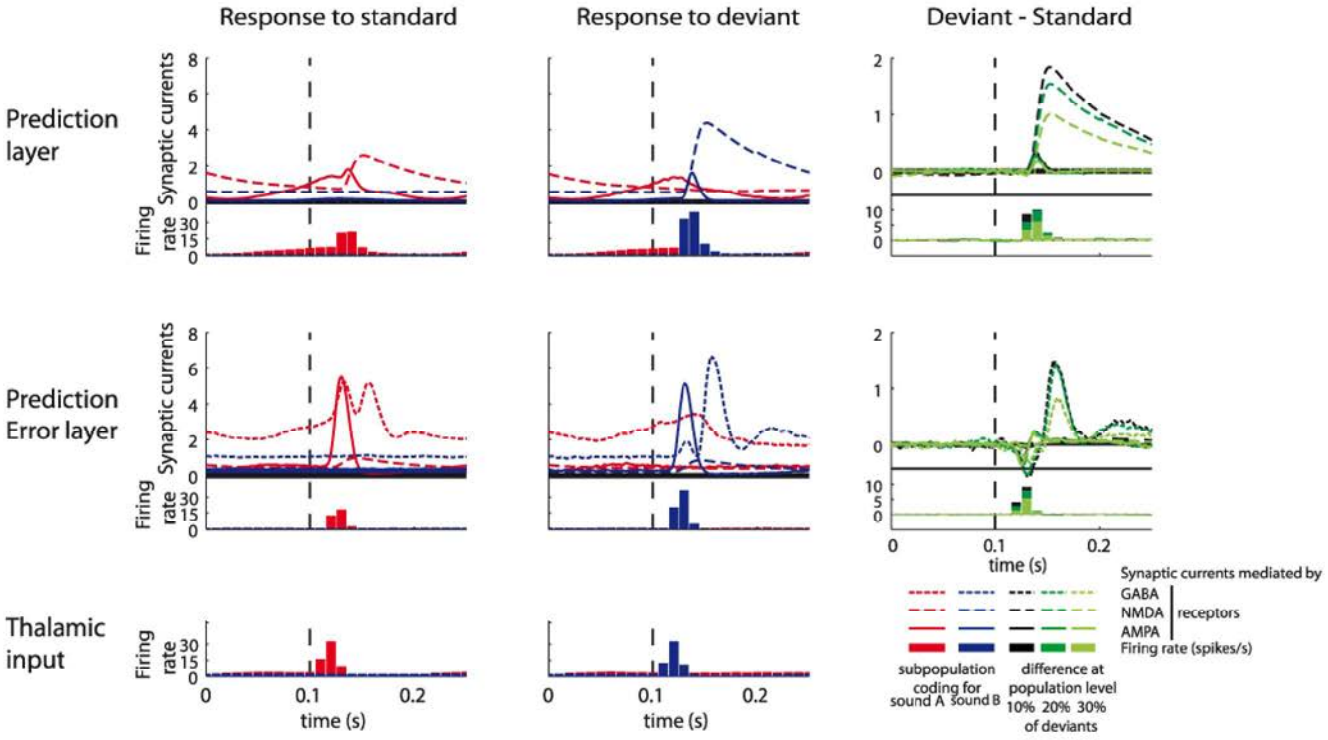


FIGURE 3.2-2 SIMULATING THE MMN IN AN ODDBALL PARADIGM : MEAN SYNAPTIC CURRENTS AND FIRING RATES.

The figure shows the mean simulated response to a standard tone (first column), a deviant tone (second column), and their difference (third column) after 200 learning trials in an oddball paradigm. Each line shows the response of a different layer of units in the model (organized as in Figure 3.2-1). For each layer, the top part of the plot represents the synaptic currents received by the subpopulation, separately for the different types of post-synaptic receptors that mediate these currents: AMPA (continuous line), NMDA (dashed line) or GABA (dotted line). The lower part of each plot displays the mean firing rate of each subpopulation. In the first and second columns, subpopulations responding to the frequent A sound (90% of trials) are represented in red, and those responding to the rare B sound (10%) in blue. The third column shows the results of simulations where the percent of deviants was varied (10, 20 or 30%).

The MMN is the result of a subtraction of the event-related potentials (ERPs) to standard and deviant stimuli. The ERPs are believed to be the result of a weighted integration of post-synaptic currents. As a simplified proxy for local field potentials or EEG responses, we calculated the difference in the sum of currents received by each layer for standard or deviant sound. Fig 2C shows the result of that operation. We can observe that there is indeed a difference in the currents between the two stimuli. For convenience, we will call this analog of the experimental phenomenon the simulated MMN or sMMN.

3.2.4.2 Behavior of the memory neurons

The memory neurons play an important role in the model. The stimulation of the network results in the activation of the predictive population either because the incoming stimulus is predicted or because of the transmission of prediction error. When the predictive population is active, it triggers the set of delay-line neurons (see Figure 3.2-3). The activity propagates linearly in the population, such that there is a direct relation between the indices of the neurons in the delay line and the temporal information coded by their activity. The precision of timing changes as a function of the interval coded: the jitter in the exact time of activation of the neurons increases with the delay coded (approximating Weber's law). Essentially, the activity of a neuron in a delay line codes for two properties of past inputs: the identity of a past stimulus and the time elapsed since the occurrence of that stimulus. The particular choice we made for the implementation of this double function (delay lines) is not fully physiologically realistic but was made for the sake of clarity and computational economy (see Discussion).

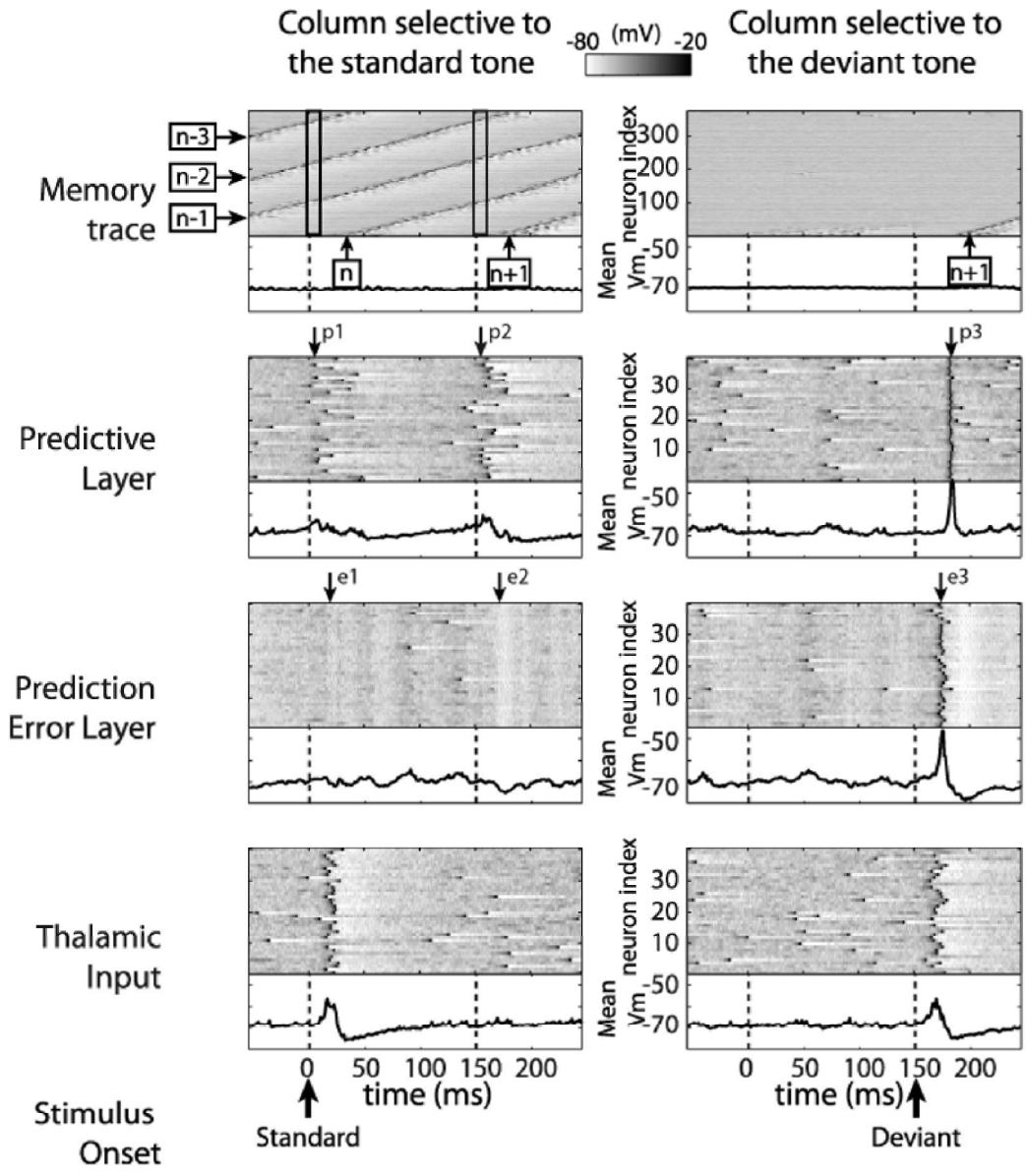


FIGURE 3.2-3: SIMULATED PATTERN OF NEURAL FIRING AND MEMBRANE VOLTAGE DURING A SINGLE TRIAL OF THE ODDBALL PARADIGM.

The figure shows a typical response to a standard tone (t=0ms) followed by a deviant tone (t=150ms). Left column, subpopulations selective to tone A; right column, subpopulations selective to tone B. For each layer, the upper part of the panel represents single-unit membrane voltage (one line per simulated neuron); the lower part is the average voltage over the population. The neurons of the memory trace are reordered so that the propagation of the activity in a synfire chain way is made obvious. "n-1"," n-2 "and "n-3" arrowed boxes refer to past stimuli whose activity is propagating in the delay lines initiated. In the left column, "n" and "n+1" arrowed boxes point to the initiation of a new memory trace following synchronous activity of the predictive population corresponding to the prediction of the stimuli n and n+1 ("p1" and "p2" arrows). In the right column, the "n+1" arrowed box shows the initiation of a new memory trace following synchronous activity of the predictive population corresponding to the prediction error signal of the n+1 (deviant) stimulus. After learning (see Figure 3.2-4), a reproducible pattern of activation in memory trace produces a depolarization in the predictive layer (black arrows) via a population of interneurons (not displayed here). The activity in predictive layer induces an hyperpolarization in the prediction-error layer ("e2" arrow) at the approximate time when an A sound is expected. At t=0 both prediction and input belong to the same column, resulting in a

cancellation of excitation and inhibition inside the prediction-error layer ("e1" arrow). At t = 150 ms, when a deviant stimulus B is presented, a depolarization of the prediction error population selective to the deviant ("e3" arrow) can be observed in parallel to the hyperpolarization of the predictive population selective to the standard ("e2" arrow). This depolarization is transmitted to the predictive ("p3" arrow) and memory (left column "n+1" arrow) populations.

3.2.4.3 Layer distribution of current sources.

We proposed a tentative localization for each functional population within the cortical layers, according to which prediction error populations correspond to granular layer and predictive populations belong to supragranular layer. Javitt et al.(1996) provided relevant intracortical local field potential data on the cortical origins of the MMN in primates. They showed in particular that the MMN mainly originates from supragranular layers of the cortex. The results of our simulations are consistent with these data, as they show that the sMMN primarly originates from synaptic currents impinging upon prediction neurons (and arising from prediction-error neurons). Importantly, note that even though there is a major difference in the firing rate of the prediction error population between the two stimuli, it does not involve a difference in the sum of synaptic inputs received by this layer as a whole, but rather a different distribution of these inputs on neurons coding for sounds A and B.

Studies in mice (Ehrlichman, Maxwell, Majumdar, & Siegel, 2008), rats (D Tikhonravov et al., 2010; Dmitry Tikhonravov et al., 2008) and monkeys (Javitt et al., 1996) also showed that MMN is strongly affected by NMDAr inhibitors. In our simulations, the sMMN results essentially from NMDAr dependent currents, which is consistent with this observation.

3.2.4.4 Effect of Deviant probability.

The vast literature on the MMN describes a broad set of properties (for review see e.g. Näätänen et al., (2007)). To evaluate the range of validity of this model, we next simulated the response of the model in various conditions mimicking classical experimental paradigms. Our first test concerned the effect of the proportion of deviants in the standard oddball paradigm. Sato et al. (Sato et al., 2000) described a systematic and parametric dependency of MMN amplitude on the probability of occurrence of a deviant sound. They showed that amplitude of the MMN increases as the frequency of the deviants decreases. We simulated the network for various proportions of deviant in the oddball paradigm (10%, 20%, and 30%). Results are plotted in Figure 3.2-2.C. We can see that the amplitude of sMMN indeed increases with the rarity of the deviants. This reduction in sMMN comes from the increased activity of the predictive population coding for B, as a result of its more frequent occurrence after an A, combined with a slightly

lower prediction of the A sound that increases the average prediction error to A. This finding closely matches the experimentally recorded ERP data .

The frequency effect shows that MMN is not an all-or-none phenomenon, but a graded response whose amplitude reflects a parametric quantification of the amount of surprise conveyed by the stimulus, given the past stimuli. It is consistent with an internal model that takes into account statistical regularities.

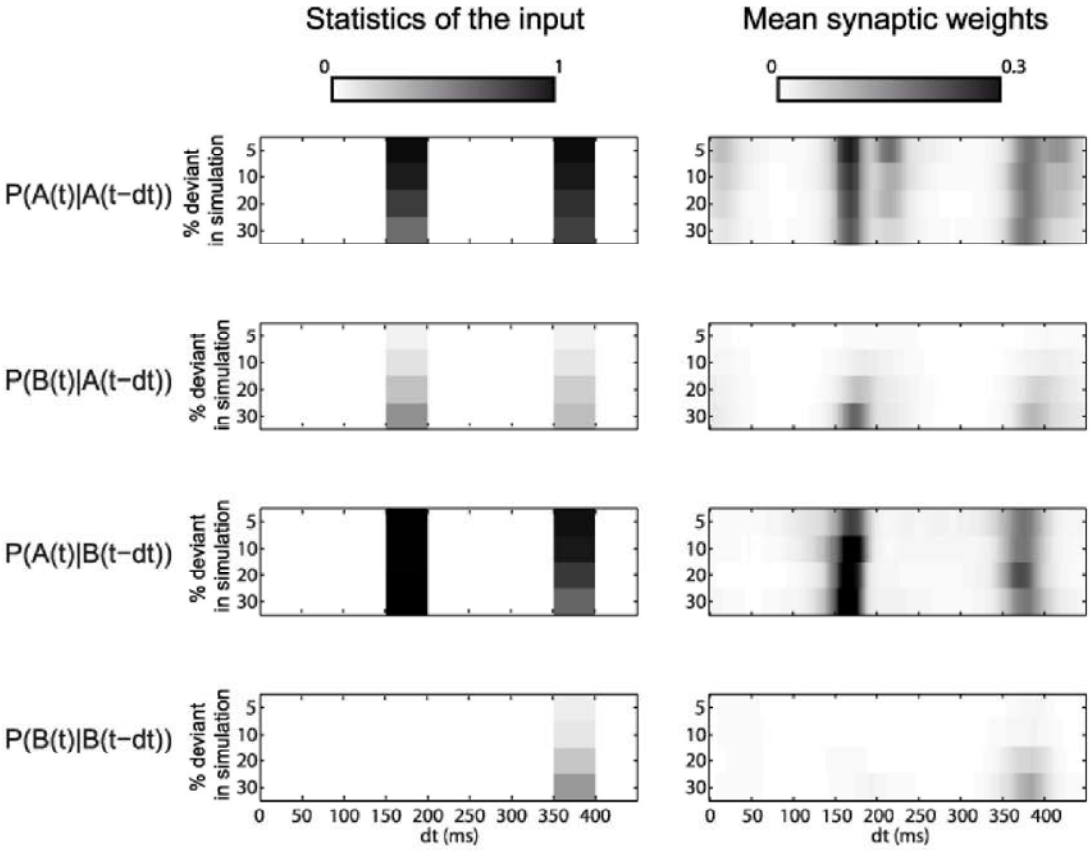


FIGURE 3.2-4: CORRESPONDENCE BETWEEN THE TRANSITION STATISTICS OF THE INPUTS (LEFT) AND THE SYNAPTIC WEIGHTS LEARNED BY THE MODEL (RIGHT).

In each pannel, the statistics are given for simulations with 5%, 10%, 20% and 30% of deviant sounds B in an oddball paradigm. Left column, conditional probabilities of receiving a given sound (A or B) at time t, given the recent history of past inputs at times t-dt (dt ranging from 0 to 400 ms). Right column, corresponding synaptic weights in the simulation at the end of learning. Gray levels indicate the mean synaptic weights between neurons of the recurrent memory network spiking on average at the time dt after the occurrence of an A or B sound, and the predictive neurons coding for the arrival of an A or B sound.

3.2.4.5 Internal model of the temporal statistics in the input

The simplicity of the population of memory neurons used in our model allows us to visualize the statistical information learned by the network (Figure 3.2-4). The only plasticity in the model occurs at synapses between the memory neurons and the predictive subpopulations.

The information coded in these synaptic weights can be directly compared to the actual conditional probabilities in the actual input sequences. Figure 3.2-4 shows the mean synaptic weights between the delay lines and the predictive sub populations as a function of the probability of occurrence of a deviant. They are compared to the actual statistics of transition probabilities in the inputs. Even though the plasticity rule was not specifically designed to converge onto a conditional transition probability, we can observe a close correspondence between the learned synaptic weights and the conditional information contained in the input. The peaks of synaptic strength coincide with the temporal intervals between the stimuli, and their amplitude is proportional to the probability of a transition between two stimuli almost regardless of the probability of occurrence of the first stimulus. Thus, this observation provides a very simple picture of what our model does: it stores, within its synaptic strengths, the conditional probability of observing a second stimulus at a certain latency after the first. Our claim is that the MMN reflects, in a quantitative manner, the degree of violation of such transition probabilities.

Importantly, the present model relies on STDP plasticity to internalize the statistics of the input. Data show that the MMN develops rapidly within few presentations of the standards (Winkler, Cowan, Csépe, Czigler, & Näätänen, 1996). To account for the MMN with such a mechanism, it is critical that plasticity occurs on a short time scale of a few seconds. To our knowledge, there is no data testing this prediction by trying to induce STDP on short time scales using ecological stimulation, and this hypothesis is therefore a prediction of the model that remains to be tested experimentally.

The time span over which the stimulus transitions can be learned is strictly limited by the capacity of the memory. Here, we adopted as a simplifying assumption the hypothesis that the memory trace abruptly vanishes after 400ms. In spite of this artificially abrupt transition, we observe that synaptic weights get progressively weaker for more distant delays, due to the increased jitter in the coding of increasingly longer temporal intervals. In a more realistic memory network, the artificial delay lines that we used could be replaced by more realistic chaotic temporal dynamics, as in "reservoir" or echo state networks models (Buonomano & Laje, 2010; Buonomano, 2005; Maass, Natschlager, Markram, & Natschläger, 2002; Pascanu & Jaeger, 2010). The memory trace would then become increasingly diluted with elapsed time, thus explaining that, in the standard oddball paradigm, a partially preserved but increasingly reduced MMN is observed as the time interval between tones is increased (Pegado et al., 2010).

3.2.4.6 MMN to repetition in an alternate signal

To further assess the properties of the model, we simulated the response to sequences where two stimuli are presented in an alternate fashion (ABABA…). On rare occasions, sound B is replaced by sound A. Horváth and Winkler (2004) showed experimentally that, in this condition, a MMN is now observed to the unexpected *repetition* of a stimulus B, in a context in which an alternation (ABABA…) was expected. This result is counter-intuitive for habituation models, but entirely compatible with predictive-coding models. Indeed, we simulated the response of the network for an input constituted by a regular alternation of A and B every 150ms. Rarely, sound B was replaced by sound A, resulting in the succession of 3 A's in a row. Results are plotted in Figure 3.2-5. An sMMN is observed, showing that the unexpected repeated sound behaves as a deviant in the standard oddball paradigm. Indeed, the predictive population coding for B increases its activity 150ms after an A occurred. In other words, the network learns to predict that after an A comes a B at 150ms. This internalization of input statistics can also be seen in the synaptic weights.

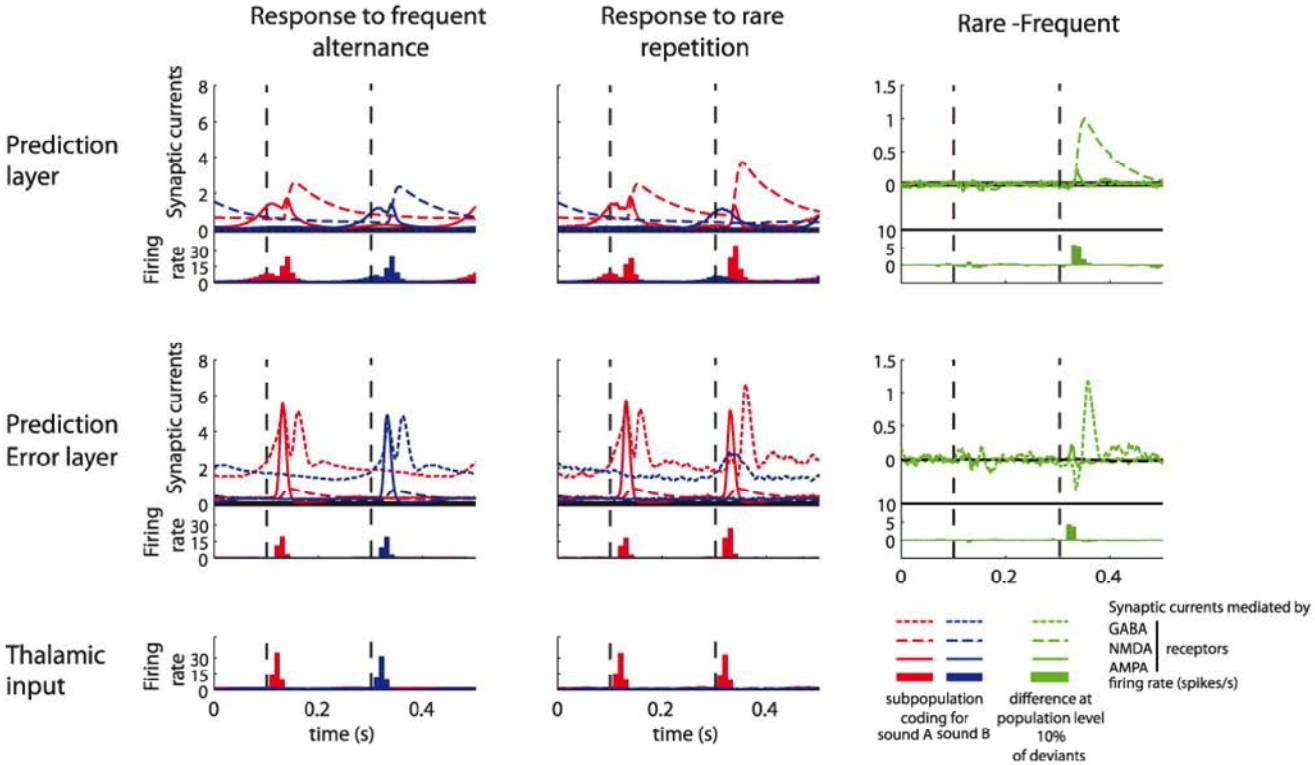


FIGURE 3.2-5: SIMULATING THE MMN IN RESPONSE TO AN UNEXPECTED REPETITION AMONGST ALTERNATING STIMULI.

Left column : mean response of the model to a frequent AB alternation in a ABABABA… stimulus. Middle column : mean response to the rare AA repetition. Right column. The difference between the rare repetition and the frequent alternation shows a MMN elicited by the repeated sound AA. This prediction distinguishes predictive coding models

3.2.4.7 Blindness to global regularities

Experimentally, the MMN is known to be blind to some global regularities in the stimulus sequence. For example, Bekinschtein et al. (2009) showed that, when participants are presented with the repetition of a 5-tone sequence AAAAB, the final B sound continues to elicit a MMN even though the occurrence of this sound is perfectly predictable based on the prior occurrence of four A sounds. In other words, the MMN seems to be "blind" to the overall sequence, and sensitive primarily to local transition probabilities, which favor the A→A transition over the A→B transition. Figure 3.2-6 shows the result of the simulation of our network on this paradigm. 150 sequences of five inputs with ISI of 150ms were presented. 70% were AAAAB sequences 20% AAAAA and 10% AAAA (omission of the last sound, not analyzed here). The SOA between two sequences was 1.2s. The average response to a frequent sequence is plotted in Figure 3.2-6. Note first that the first element of the sequence is not predicted. The time elapsed since the last sound is superior to the span of the delay line. It is consistent with data showing that no MMN exists on the first element of a sequence or for very long ISI (N. Cowan et al., 1993; Mäntysalo & Näätänen, 1987). Second, the final B sound elicits a stronger prediction error (sMMN) than the previous sounds. This effect arise because (1) the transition probabilities favor the prediction of an A sounds following an A sound; and (2) the network cannot use the past occurrence of a B sound to predict a new B sound, because the temporal interval between them (1200ms) exceeds the time span of the memory neurons. Both the increased response to the first sound and the final MMN tightly reproduced experimental scalp and intracranial recordings (Bekinschtein et al., 2009; Wacongne et al., 2011)

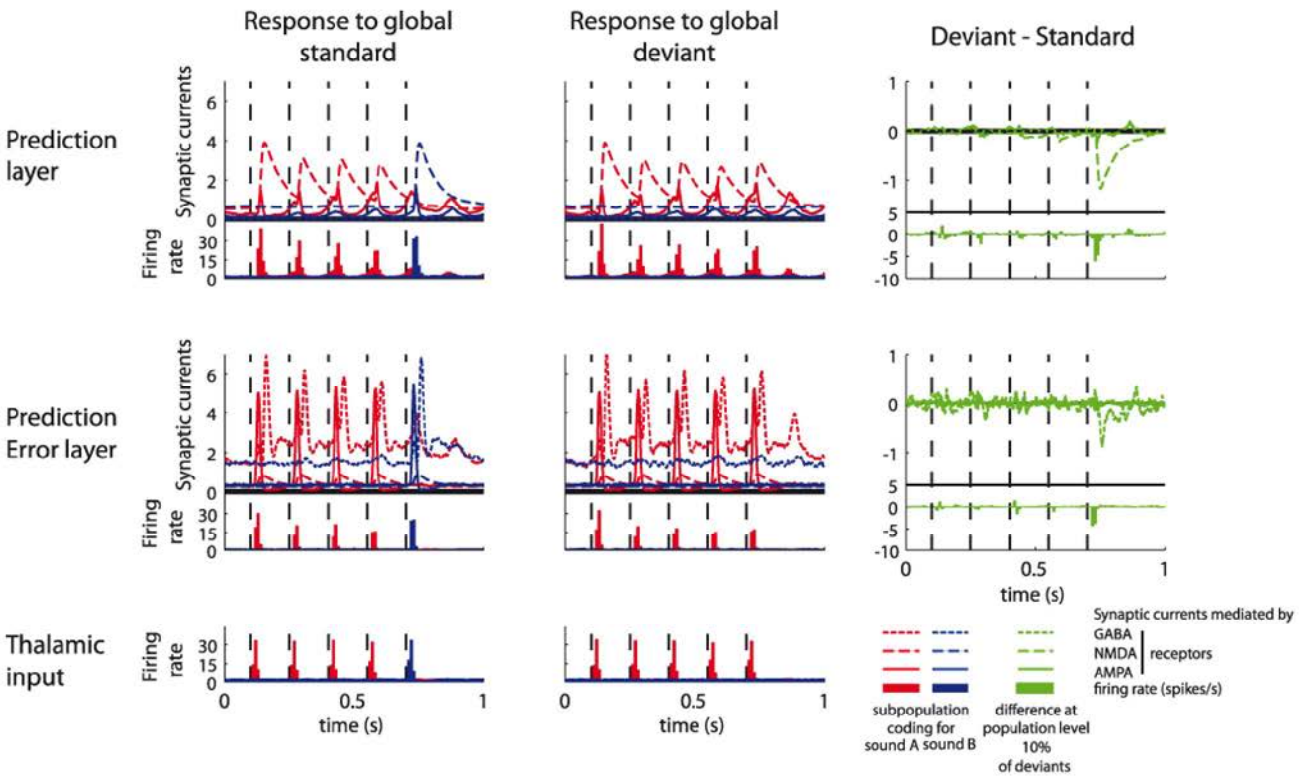


FIGURE 3.2-6: SIMULATING THE LACK OF SENSITIVITY OF THE MMN TO GLOBAL REGULARITIES THAT CANNOT BE CAPTURED BY LOCAL TRANSITION STATISTICS.

Left column : mean response to a frequent AAAAB stimulus. Middle column : mean response to the rare AAAAA stimulus. Right column : difference between the rare repetition and the frequent alternation. A MMN continues to be elicited by the final B sound of the standard AAAAB stimulus. Although the global sequence AAAAB is frequent and predictable, the MMN effect is driven primarily by the rarity of the local transition A→B.

Using a closely related, yet importantly different paradigm, Sussman et al. (1998) showed that the MMN to the deviant sound B in circular sequences AAAABAAAAB… actually disappears if the SOA is small (100 ms) and B is presented at regular intervals. This observation is actually consistent with the model we propose. If the time between two B sounds is short enough, the network is able to learn the transition between two consecutive Bs, and the sMMN disappears. Our simulated network predicts that the MMN should reappear as soon as the temporal prediction of B is made impossible, either by spacing the B presentations beyond the capacity of the memory neurons, or by making B appear at irregular time intervals.

3.2.4.8 MMN to omission

One of the most remarkable properties of the auditory system is that it can generate evoked responses to an absent but expected stimulus (Hughes et al., 2001; Joutsiniemi & Hari, 1989; Raij et al., 1997; Wacongne et al., 2011; Yabe et al., 1997). We similarly tested the response of our network to the omission of an expected sound. We simulated the response of the network

to pairs of AB sounds (ISI = 150ms) separated by 500ms, and rarely (10% of trials) omitted the second tone of the pair. We compared the response to such omissions to the response to identical single A tones presented every 500ms in a block where they were the only stimulus, and therefore no second stimulus was expected. As shown in Figure 3.2-7, the predictive currents anticipated the arrival of a second B sound and therefore produced a response to a non-existing sound, as experimentally observed. Indeed, our results are tightly consistent with intracranial data obtained on a similar protocol (Hughes et al., 2001).

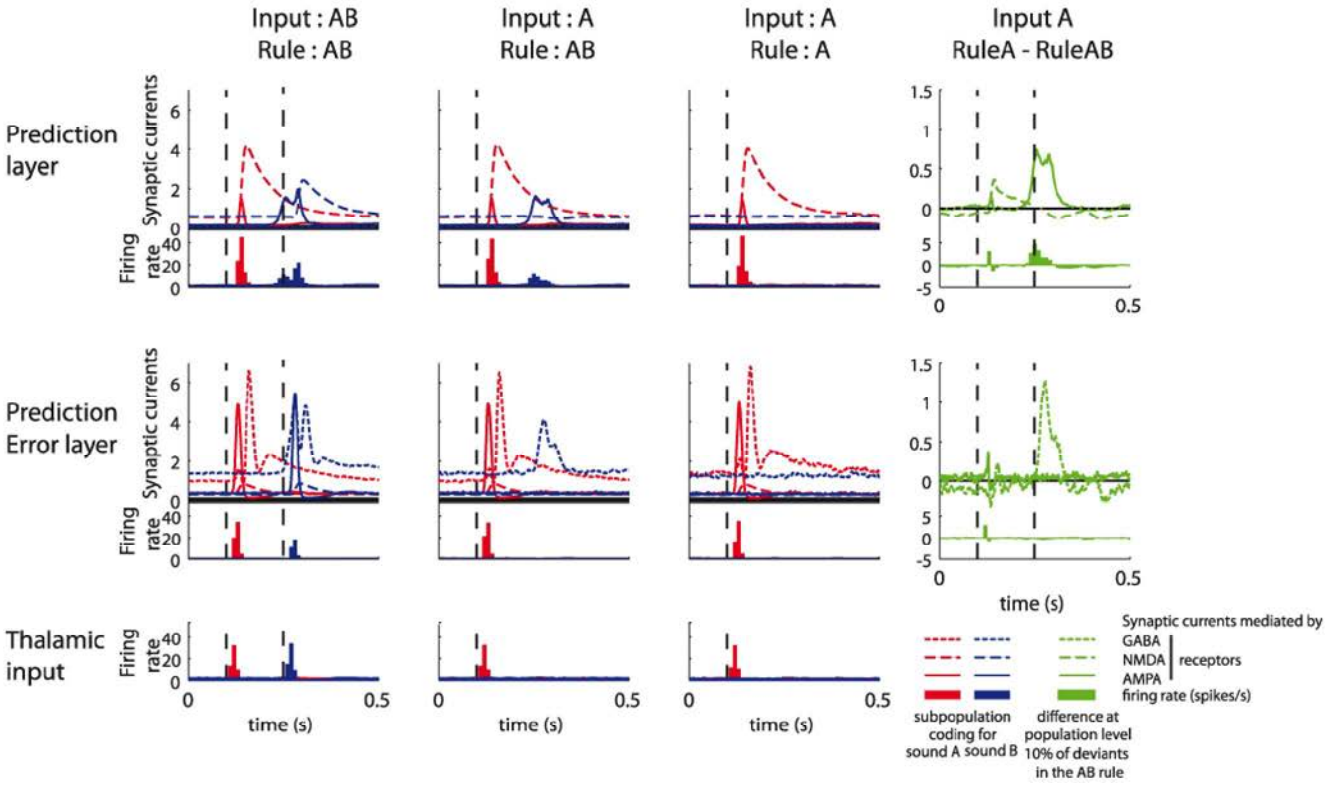


FIGURE 3.2-7: SIMULATING THE MMN TO THE OMISSION OF AN EXPECTED SOUND.

First column : mean response to a frequent AB pair. The network learns the predictable local transition A→B, which results in a reduced response to the predictable B sound (see arrow). Second column : mean response to a rare A sound presented in isolation in the same context. The network generates a response to the omission of the expected sound B (arrow). Third column : response to the same isolated sound A, in a different context where it is the frequent stimulus. Although the stimulus is physically identical to the second column, the predictive response to the omitted B sound is no longer seen. Fourth column: difference between the second and third columns, isolating the simulated MMN to omission.

Interestingly, although this omission response is frequently called a MMN in the literature, our model proposes that it does not have exactly the same computational significance as the classic oddball MMN. In a predictive coding model, the omission response reflects solely a predictive component and not a prediction error *per se*, i.e. it does not reflect late, NMDA

dependent, prediction error currents, but early predictive currents. In the oddball paradigm, the main origin of the difference is a NMDA dependent supragranular current, whereas the model predicts that the omission response should be resistant to competitive antagonist of NMDA channels, once the transition probabilities are learned.

3.2.4.9 MMN to changes in duration

Until now, we only simulated the onset of the input sounds. However, in primary auditory cortex, there are also populations of neurons that respond to sound offset (Chimoto, Kitama, Qin, Sakayori, & Sato, 2002; Volkov & Galazjuk, 1991). In a predictive coding perspective, the mechanism that we describe should capture not only how the onset of one sound can be predicted from the onset of another, but also how the offset of one sound can be predicted based on the onset of the *same* sound. In the present section, we show that this effect can explain the observation of a MMN to a change in sound duration.

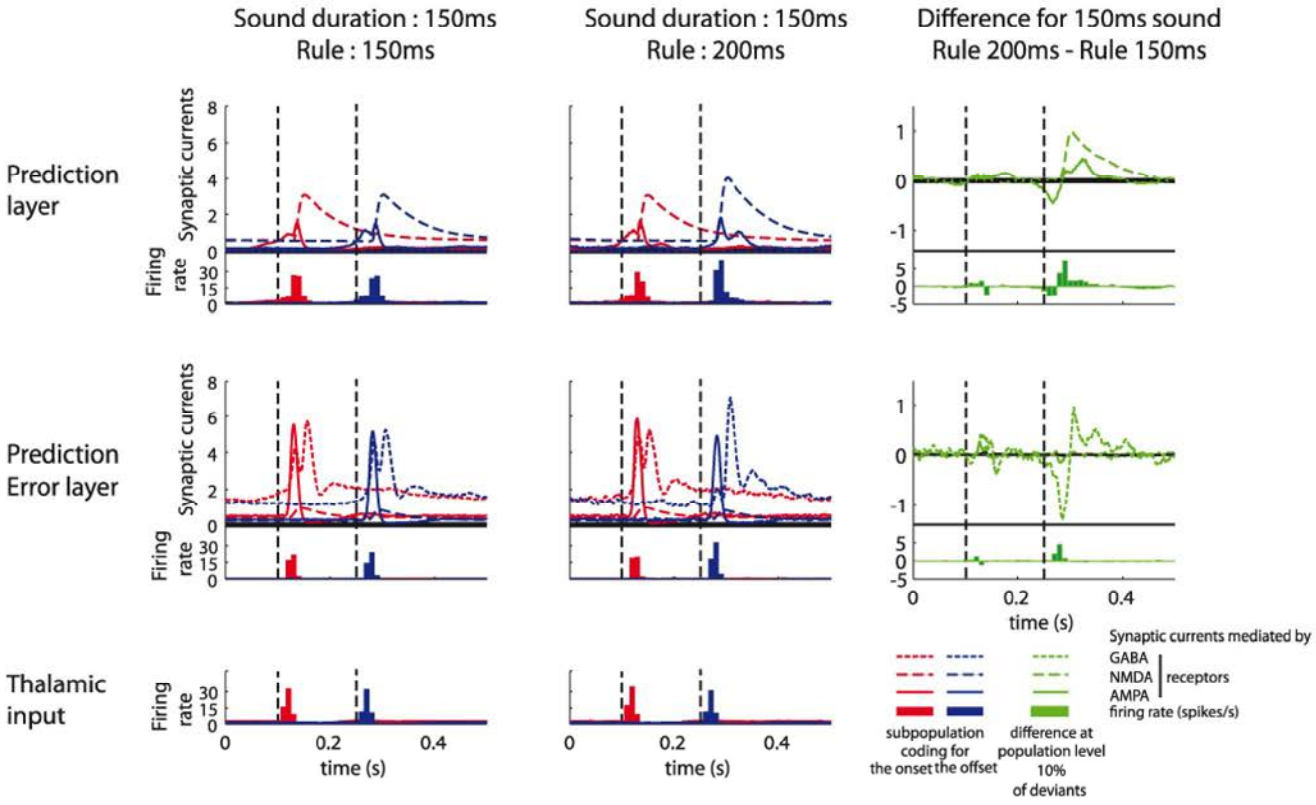


FIGURE 3.2-8: SIMULATING THE MMN TO A DURATION DEVIANT.

Blue and red now represent subpopulations selectively responsive, respectively, to sound onset and offset. Left column : response to a frequent 150 ms long sound. Middle column : response to the same physical 150 ms sound when it serves as the rare deviant in an oddball paradigm where the frequent sound is 200 ms long. Right column : difference between these two responses, isolating the MMN evoked by an unexpected change in duration.

We stimulated our network with sounds of 150ms duration, separated by a 300ms ISI. We now assumed that the neural population "A" responded to the onset of the stimulus, and the "B" population to the offset. On a rare 10% of trials, the duration of the sound, that is the interval between the onset and the offset of the sound, was changed to 200ms. We also simulated the converse situations where standard sounds were 200ms long and deviants, 150 ms long. Results are plotted in Figure 3.2-8, where we compare the response to two physically identical sounds (150ms duration) that act as standards or as deviants. When the input duration deviates from expectations, the internal model generates a prediction later than the actual arrival of the stimulus. The response to the offset is not cancelled and the prediction error is bigger. This prediction-error signal is followed by another component, corresponding to the response to the omission of the later onset. Altogether, these responses capture the experimentally observed MMN to duration deviants (Jacobsen & Schröger, 2003).

Note that, in our model, the change in duration is formally equivalent to a change in interstimulus interval (ISI): predictions that are focused in time fail to cancel incoming inputs that are shifted in time. Therefore the model also reproduces the experimentally observed MMN to ISI deviants (Ford and Hillyard, 1981; Nordby et al., 1988).

3.2.4.10 Prediction vs. habituation : an experimental test of the model

We have shown that a model exclusively based on predictive coding principles can explain, on a parsimonious basis, the major properties of the experimentally observed MMN. However, this is not the only theory proposed in the literature. May and Tiitinen (2009) defend the theory that MMN would only be the result of synaptic habituation, that is to say, the reduction of the amplitude of EPSPs as a result of repeated stimulation of the same synapse. Indeed, synaptic adaptation and short term plasticity are commonly observed in vivo and in vitro in cortex (see Calford, 2002, for review), and more specifically in auditory cortex (see e.g. Condon and Weinberger, 1991; Brosch and Schreiner, 2000) and it is likely that a complete theory of MMN should ultimately take such effects into account. However, is synaptic habituation sufficient to explain all MMN findings? In their review of MMN findings, May and Tiitinen (2009) suggest that all current MMN paradigms remain compatible with a habituation mechanism, and argue that there is therefore no decisive evidence in favor of predictive coding models of the MMN.. Contrariwise, our model leads us to propose one such critical test separating the predictive coding and habituation interpretations.

To provide a direct test of the two models, we decided to present pairs of closely consecutive sounds AB (200 ms SOA), separated by a broad temporal interval (>10 seconds). Occasionally, instead of the frequent AB pair (70% of trial), a deviant AA pair is presented in 10% of the trials, in which the same sound is repeated twice. The predictions of our model are straightforward: the first A sound predicts the second B sound in the frequent AB pair, and a mismatch negativity should therefore be generated whenever the unexpected A sound is heard instead (i.e. when the rare AA pair is presented instead of the frequent AB pair). We confirmed this prediction through simulations (the results are essentially identical to the alternation case ABABA… described earlier).

The habituation model, however, makes the opposite prediction: due to synaptic habituation, the second A sound in the AA pair should always elicit a reduced activity compared to the B sound in the AB pair, which solicits non-habituated synapses. It could be argued that some higher-order neurons might habituate to the presentation of the frequent AB pair as a whole. Indeed, this is how May and Tiitinen, (2009) account for the above-describe alternation paradigm (ABABA…). However, experimentally, the recovery time of synaptic depression is generally of the order of a few seconds (Ulanovsky et al., 2004; Varela et al., 1997). Thus, by making the temporal interval between pairs as long as 10 seconds, we should render this putative effect of synaptic habituation at the level of the whole pair quite negligible, especially as compared to the short-term adaptation to the individual sounds A in the pair AA, which are only separated by 200ms. In this case, the habituation model can only predict a reduced brain response to the infrequent AA pair, i.e. the converse of a mismatch negativity.

Legend of the figure 3.2-9. (A) Experimental design. Each block of trials begins with 10 identical pairs of tones (A followed by B). A and B are pure tones of 50 ms and frequency 800Hz and 1600Hz, counterballanced between blocks and subjects. The subject then listened to 120 pairs of tones: 70% of frequent AB pairs, and 10% of each of the rare pairs AA, BA, and BB. (B) Comparison beween the relative response amplitude predicted by the habituation model, the predictive coding model, and the data. In the habituation model (left column), response amplitude is minimal to a repeated tone. In our predictive coding model (middle column) response amplitude depends on transition probabilities between the first and second tone of the pair. The two models generate qualitatively different prediction for the AB and AA pairs. Observed group level responses (right column) to the two tones of each pair fit with predictive-coding predictions (see methods for details). Error bars represent the standard error of the mean. (C-F) MEG results for magnetometers for one representative subject (left) and for the average over all subjects (right). Panels C and E show the sensor-level topography of the average difference in magnetic field between the rare AA and the frequent AB pairs, 170ms after the onset of the second sound. The most significant cluster of sensors at this time is indicated by dots. Panels D and F show the time course of the average response to all conditions within these sensors. Line colors correspond to the brackets surrounding the stimuli in panel A. The black line above the curves indicates the interval where a significant difference was found between AA and AB.

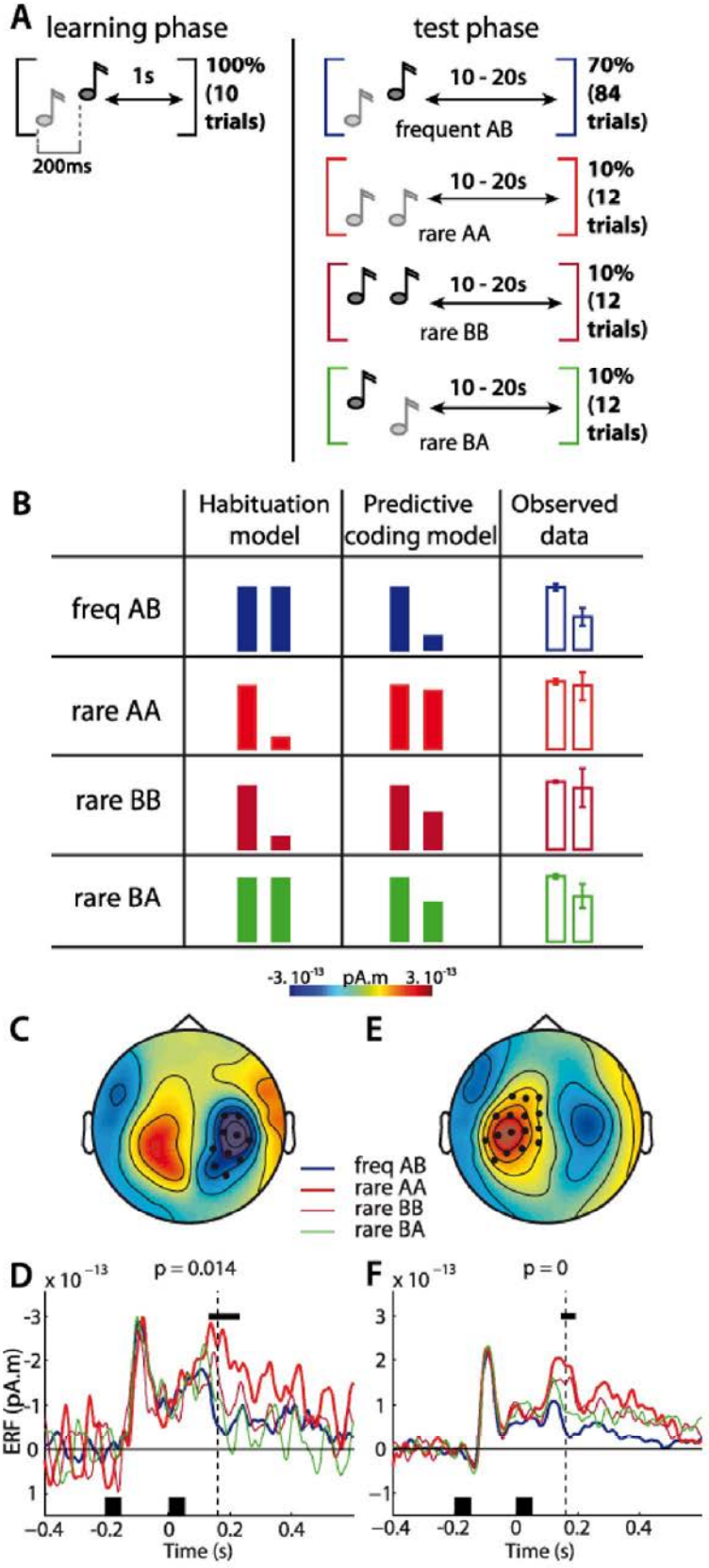


FIGURE 3.2-9: EXPERIMENTAL TEST OF THE MODEL USING MAGNETO-ENCEPHALOGRAPHY.

As a further control, we introduced two additional rare deviants, the BB and BA pairs, which were also presented in 10% of the trials each. These pairs have the same structure as the AA pairs and AB pairs, but are presented with equal probability. In our model, as the transition probabilities B→B and B→A are the same, the predicted evoked responses should be the same. Thus, our model predicts a lack of any difference here, where as the synaptic habituation model again predicts a reduced response to the repeated pair BB compared to the non-repeated pair BA

We recorded MEG signals while 5 healthy participants were instructed to listen to these stimuli. Each subject listened to two blocs of 120 pairs of sounds. The frequencies of the two sounds were 800Hz and 1600Hz, and were counterbalanced between blocs. Figure 3.2-9 shows the results. In every subject, the second tone of the rare AA pairs elicited a MMN compared to the frequent AB pairs. The difference between the two conditions was significant for each individual subject and for both types of sensors (sujet1: Grad : 121-206ms, $p < 1e-16$; Mag : 131-231ms; sujet2 : Grad, 131-186ms, $p = 0.028$; Mag, 157-204ms, $p = 0.044$; sujet3 : Grad, 127-226ms, $p = 0.003$; Mag, 126-264ms, $p = 0.004$; sujet4 : Grad, 109-177ms, $p = 0.006$; Mag, 110-230ms, $p = 0.001$; sujet5 : Grad, 120-164ms, $p = 0.04$; Mag, 116-260ms, $p = 0.01$), as well as at the group level (Grad, 108-232ms, $p<1e-16$; Mag, 145-193ms, $p<1e-16$). The topography of the effect was similar to the classical MEG-MMN topography, with bilateral temporal activations.

Our model predicted that no difference should exist between the two control stimuli BA and BB. Indeed, no significant difference was observed between the two control stimuli (rare BB and rare BA pairs, presented with equal probability). In fact, a non-significant trend existed in the direction opposite to the one predicted by the synaptic habituation model (greater brain response to BA). This finding can be explained by the fact that the identity of the sounds serving as A and B was counterbalanced between the two halves of the experiment. As a result, the rare BA pair of the second run was the frequent AB pair of the first run. We reasoned that the transition that was well learned during the first block of trials could have continued to prevail in the second block, especially as the pairs BB and BA were presented for a very small number of times (12 each), thus largely preventing relearning of the actual equiprobability of the B→A and B→B transitions. We confirmed this hunch by separately analyzing the first and second halves of our experiment. When restricted to the first half, the two control stimuli BA and BB did not present any identifiable difference, whereas the same two conditions presented a stronger (yet non-significant) difference in the second half. Note again that the latter difference (stronger response to BB) was in the direction opposite to that expected from a habituation mechanism.

The experimental data are therefore consistent with the predictions of our model in great detail and in every single subject. To explain the data with synaptic habituation, one would have to postulate the existence of neurons that (1) respond specifically to the transition between the AB sounds; (2) present significant habituation after 10s; and (3) whose habituation to AB pairs is strong enough to override the counter-effect of habituation to the AA pair for neurons that respond only to frequency A. The latter assumption is particularly implausible because neurons responsive to A alone are likely to be much more numerous than neurons responsive to the AB pair as a whole, and that their habituation would be likely to be much stronger, given that the A-A delay of 200 ms is much shorter than the AB-AB delay of 10s or more. Furthermore, the responses to BA and BB pairs provide no support for a habituation to individual B sounds. We therefore conclude that any habituation account of our data seems highly implausible.

3.2.5 Discussion

In this study, we developed a spiking neuron model of mismatch negativity, based on a predictive coding approach. We identified key properties of the mismatch effect and simulated the network response to a variety of test sequences. In particular, our model reproduced the known reduction in MMN amplitude when the frequency of the deviants increases, the MMN to repetition in an alternate sequence, and the response to the omission of an expected sound. Without any additional assumption, the model was able to account for the MMN to a change in stimulus duration or in ISI. We proposed a precise cortical localization of the neuronal populations postulated in the model and showed that our simulated currents sources were consistent with actual electrophysiological data. We also showed that the model acquired a quantitative synaptic representation of transition probabilities. An alternative model hypothesizes that MMN arises purely from synaptic habituation. We identified a precise experimental context where the two models lead to opposite predictions, and showed that MEG data from human participants fully support our predictions, with no evidence of a synaptic habituation effect.

3.2.5.1 Predictions versus synaptic habituation

In the present study, we showed that a model based on pure predictive coding, without any synaptic habituation component, could account for a large range of effects. It is important to note that even though the habituation and predictive/memory accounts of MMN have been often opposed (May & Tiitinen, 2009; Näätänen et al., 2005; Winkler, 2007), the two hypotheses are not logically exclusive. It remains possible that the two processes concur to the final MMN effect, possibly in different proportions according to the paradigm. However, the conclusions of

the MEG experimental test of our model are fully consistent with a purely predictive account of MMN, and argue against a strong contribution of habituation effects.

Other recent studies argue in favor of a negligible role of habituation in the MMN effect. Recent human MEEG recordings indicate that the omission response observed when an expected sound fails to occur conforms to the predictions of hierarchical predictive coding models (Wacongne et al., 2011). In rodents, Farley et al., (2010) showed that stimulus specific adaptation is indeed observed in auditory cortex but that its properties differ sharply from those of the MMN, in terms of sensitivity to NMDA antagonists or elicitation of a novelty response. Taken together, these results provide strong evidence against a predominant role of synaptic habituation in the MMN effect, and argue for the predictive coding hypothesis. Similar conclusions have been recently reached by other groups (Todorovic, van Ede, Maris, & de Lange, 2011).

3.2.5.2 Extensions and limits of the model

In this study, we limited our simulations to two cortical columns coding for features distinct enough that thalamic inputs did not stimulate both columns at the same time. The model could be easily extended to a more continuous coding of tone frequency, where each neuronal population codes for one preferred frequency but also responds more weakly to neighboring frequencies. This would give an account of the increase of MMN amplitude with the difference in frequency between standards and deviants (Sams, Paavilainen, Alho, & Näätänen, 1985) .

Predictive coding requires that a memory of the recent past be used to predict the future. For the sake of simplicity, we adopted here the simplest hypothesis for a neural memory: a delay line. Although this assumption may not seem very realistic, we only argue here that there must be neural populations whose activity contains information about both the identity of recent stimuli and the time elapsed since they occurred. As noted by Buonomano (2005), these neurons need not be ordered in cortical space, but could be intermixed and arise from the partially chaotic temporal dynamics of cortical activation spread. Electrophysiological recordings from auditory cortex slices suggest that such a code might exist within the auditory cortex (Buonomano, 2003): when cortical neurons were stimulated, they triggered other neurons with reliable delays, without any correlation between response delays and the cortical distance from the neuron initially stimulated. Such a code would be ideal to support a memory of the recent past, as required in our model. It would allow the same neuronal populations to code tonotopically for the present and non-tonopically for the past.

According to this hypothesis, our entire model would fit within a single cortical column, and could constitute a basic building block for sensory predictive learning in various sensory systems. As noted by Friston et al. (1995), the closely similar neuronal architecture of cortical layers throughout the cerebral cortex, supports the view that a similar computational principle of predictive coding may apply to the multiple hierarchical levels of brain's cortical areas. Thus, our model may be used to account for higher-order instances of mismatch responses, such as the distinct MMNs evoked by a change in phoneme versus speaker (Dehaene-Lambertz, 1997; Giard et al., 1995) , or the mismatch responses observed outside the auditory modality, either in visual (Pazo-alvarez et al., 2003; Tales et al., 1999), olfactive (Krauel et al., 1999; Pause & Krauel, 2000) and somatosensory (Kekoni et al., 1997; Shinozaki et al., 1998) modalities or even in a sensory-motor context.

Our model makes clear predictions as to the kind of regularities that should be reflected by the MMN. It is only able to predict incoming stimuli by acquiring an internal representation of the transition probabilities between their onsets and offsets, over a window of a few hundreds of milliseconds. Thus, it fails to detect deviance from a rule that cannot be described at the level of transition probabilities. This statement should help clarify the issue of whether the MMN reflects "rule-based learning", which is often confused in the present literature.

For example, Sussman et al., (1998) showed that when the oddball paradigm was slightly modified so that deviant sounds B occurred regularly at short-enough intervals between the standards (AAAABAAAABAAAAB…), the MMN disappeared. Yet in a seemingly contradictory finding, using a minimally different paradigm, Bekinschtein et al (2009) showed that an AAAAB rule could not be acquired by low-level sensory processing, since the final B sound continued to elicit a MMN even when the entire AAAAB sequence was fully predictable. According to our model, the main difference between the two protocols is the long additional temporal gap between two 5-tone sequences that exist in the Bekinschtein paradigm, and which disrupts any recent memory capable of predicting the final B sound. Thus, the apparent inconsistency in the results is easily understandable if we consider the size of the memory delay needed for temporal prediction. This example stresses the importance of carefully assessing the matrix of transition probabilities when trying to design experiments probing rule learning.

An MMN-like response was also recorded for deviance from more abstract kinds of regularities such as tone repetition or ascending/descending tones (Endress, Dehaene-Lambertz, & Mehler, 2007; Korzyukov, Winkler, Gumenyuk, & Alho, 2003; Paavilainen et al., 1999). Whether or not such rules are learnable by our network depends on the specifics of the

experimental design. To make the rule unlearnable by transition probabilities, the design should reserve a broad frequency band never presented during training, or over which the probabilities of ascending and descending tones are equal. Otherwise, given enough training exemplars, our network will learn the "rule" and even generalize to frequencies that are novel but close enough to the training frequencies. These conditions were not fulfilled in many previous papers. If they were, however, and if the MMN resisted to such a control, this would provide definitive evidence that the mechanisms underlying the MMN go beyond our basic transition-probability model. The model might be extended, however, by postulating higher-order neurons sensitive to melodic contours (e.g. any ascending contour). In general, the coding properties of the input neural populations will have a crucial impact on the kind of regularities that can be detected by our model.

3.2.6 Conclusion

The idea that the brain is not a passive input-output device, but acts as a predictive system capable of anticipating on the future, has a long history in ethology, psychology and neuroscience and has been proven useful in many distinct domains of perception, cognition and action (Stanislas Dehaene & Changeux, 1991; Hosoya et al., 2005; Schultz et al., 1997; R.S. Sutton & Barto, 1998b) . Understanding the neural mechanisms by which the brain generates predictions is therefore an important goal for neuroscience. Predictive coding models of the MMN have been previously proposed (Friston et al., 2006; Friston, 2005; Garrido, Kilner, Kiebel, & Friston, 2009; Spratling, 2010), but only as abstract mathematical descriptions without a precise neurobiological implementation (Marreiros, Kiebel, Daunizeau, Harrison, & Friston, 2009) (although see (Fiorillo, 2008)). The present model resolves the difficulties associated with a neurobiological implementation of predictive coding. We show how the subtraction of observed versus predicted signals can be implemented through a specific architecture of inhibitory interneurons. We also show that a NMDA-dependent STDP plasticity rule is well adapted for learning of stimulus associations, leading to the prediction of a precise and essential contribution of NDMA receptors to predictive coding. It could generalize much beyond the specific domain of the MMN for which it was presently tested.

3.2.7 Acknowledgement

The NeuroSpin MEG facility was sponsored by grants from INSERM, CEA, the Fondation pour la Recherche Médicale, the Bettencourt-Schueller foundation, and the Région île-de-France. This project was supported by a senior grant of the European Research Council to S.D. (NeuroConsc program), as part of a general research program on functional neuroimaging

of the human brain (Denis Le Bihan). We thank Virginie Van Wassenhove, Alain Destexhe and Karim Benchenane for useful discussions.

CHAPTER 4

ARTICLE 2: EVIDENCE FOR A HIERARCHY OF PREDICTIONS AND PREDICTION ERRORS IN HUMAN CORTEX

Published in PNAS(2011)

4.1 Introduction to the article

4.1.1 Goal of the study

In the previous chapter, I proposed a model based on predictive coding principles that was able to reproduce the main properties of the MMN. A key prediction of this model is that the response observed when an expected tone is omitted reflects purely predictive activity.

The hierarchical predictive coding framework proposes that the cortical canonical microcircuit implements a predictive coding module that is replicated at different levels of the cortical hierarchy. Predictions errors would passed on to higher levels of the cortical hierarchy where the new predictive module would attempt to cancel out residual errors. The existence of multiple predictive modules that share similar computational principles is consistent with the fact that the MMN response can be observed in response to violations of temporal regularities regarding a large number of stimuli features and in different sensory modalities. However, there is no data showing hierarchical predictive responses.

The paradigm developed by Bekinschtein et al. (2009) used sequences that presented nested level of regularities. The data showed that violation of each level of regularity elicited different responses. This study adapts this paradigm, adding trials with omission of a stimulus to test the hypothesis that multiple level of regularity do generate multiple omission responses.

4.1.2 Reference

Wacongne, C., Labyt, E., Van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. Proceedings of the National Academy of Sciences, 108(51), 20754–59. doi:10.1073/pnas.1117807108

4.2 Article

4.2.1 Abstract

According to hierarchical predictive coding models, the cortex constantly generates predictions of incoming stimuli at multiple levels of processing. Responses to auditory mismatches and omissions are interpreted as reflecting the prediction error when these predictions are violated. An alternative interpretation, however, is that neurons passively adapt to repeated stimuli. We separated these alternative interpretations by designing a hierarchical auditory novelty paradigm and recording human electro- (EEG) and magneto-encephalographic (MEG) responses to mismatching or omitted stimuli. In the crucial condition, participants listened to frequent series of four identical tones followed by a fifth different tone, which generates a mismatch response. Because this response itself is frequent and expected, the hierarchical predictive coding hypothesis suggests that it should be cancelled out by a higher-order prediction. Three consequences ensue. First, the mismatch response should be larger when it is unexpected than when it is expected. Second, a perfectly monotonic sequence of five identical tones should now elicit a higher-order novelty response. Third, omitting the fifth tone should reveal the brain's hierarchical predictions. The rationale here is that, when a deviant tone is expected, its omission represents a violation of two expectations: a local prediction of a tone plus a hierarchically higher expectation of its deviancy. Thus, such an omission should induce a greater prediction error than when a standard tone is expected. Simultaneous EEE-MEG recordings verify those predictions, and thus strongly support the predictive coding hypothesis. Higher-order predictions appear to be generated in multiple areas of frontal and associative cortices.

4.2.2 Introduction

According to the "predictive coding" hypothesis , the architecture of cortex implements a top-down prediction algorithm that constantly anticipates on incoming sensory stimuli. Each cortical area houses an internal model of the environment, which is generated by compiling the statistical regularities that govern past inputs. This model is used to generate top-down predictions that are compared to novel incoming inputs. Only the difference, called the "prediction error", is transmitted to higher cortical stages, where it can be used to adjust the internal model. Importantly, this process can be hierarchically organized (Friston, 2005; Kiebel et al., 2008; Kiebel, Kriegstein, et al., 2009; R. P. Rao & Ballard, 1999), such that the prediction error arising from a given area in turn serves as the input to the next area. The outcome is an active system that constantly updates models of its environment at multiple hierarchically organized levels.

While considerable evidence supporting predictive coding has been provided at the perceptual level (e.g. 5-10), here we specifically set out to test the notion of a *hierarchy* of predictions, using a novel variant of the classical auditory violation paradigm (Bekinschtein et al., 2009).

When a rare sound is introduced within a sequence of repeated frequent sounds, it elicits a novelty response in the event-related potential, which has been termed the mismatch negativity (MMN) (Näätänen et al., 1978). This response is interpreted, within the predictive coding framework, as reflecting the violation of a prediction: the MMN would directly reflect the cortical prediction error signal (Friston, 2005; Garrido et al., 2007; Garrido, Kilner, Kiebel, et al., 2009; Winkler, 2007). This interpretation is supported by sophisticated modeling studies which suggest that the MMN can only be accounted for by postulating a top-down predictive contribution (Garrido et al., 2007; Garrido, Kilner, Kiebel, et al., 2009; Parmentier et al., 2011; Todorovic et al., 2011). However, an alternative interpretation exists, whereby the MMN would solely reflect a passive, bottom-up process of adaptation to the repeated stimuli (Garrido, Kilner, Stephan, & Friston, 2009; May & Tiitinen, 2009). According to this interpretation, the repeated stimulus, by constantly stimulating the same afferent pathways, leads to synaptic adaptation and therefore to a reduced activation. The rare stimulus, by contrast, activates fresh afferents which have not been adapted, resulting in a distinctly larger response. Thus, mismatch responses could reflect adaptation rather than predictive coding.

How could adaptation and predictive models be distinguished? An interesting variant of the mismatch paradigms consists in omitting the expected stimulus, rather than substituting it

with another stimulus. It is a rather remarkable fact that the auditory cortex generates extensive responses locked to the absence of a predictable sound. This omission response can be detected by a variety of methods, including event-related potentials (ERPs) (Yabe et al., 1997), magneto-encephalography (MEG) (Raij et al., 1997), and intracranial recordings (Hughes et al., 2001). Omission responses fit quite naturally within the predictive coding framework: if stimulus-evoked brain activity indexes the difference between a sensory signal and its top-down prediction, then when the sensory signal is omitted, the evoked activity should reflect the pure prediction signal within the same cortical area (Bendixen, Schröger, & Winkler, 2009a; den Ouden, Friston, Daw, McIntosh, & Stephan, 2009; Todorovic et al., 2011). Omission responses seem more difficult to explain within the adaptation framework. They might reflect the automatic rebound of a cortical oscillator entrained by the rhythm of the past stimuli (May & Tiitinen, 2009), but this hypothesis meets difficulties explaining that omission responses are still present in non-rhythmic paradigms, for instance when omitting the second tone of a pair (Bendixen et al., 2009a; den Ouden et al., 2009; Hughes et al., 2001; Todorovic et al., 2011).

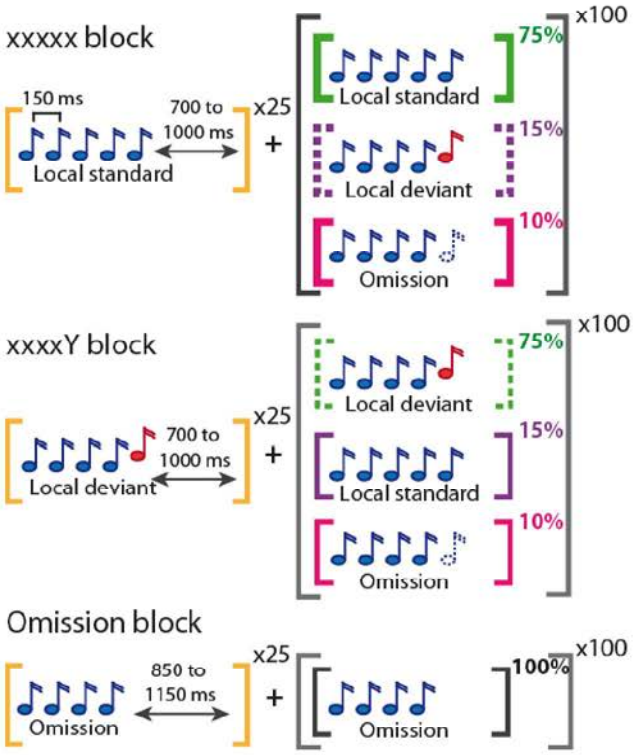


FIGURE 4.2-1 : EXPERIMENTAL DESIGN.

Three auditory stimuli could be presented: local standards (a series of five identical tones, denoted xxxxx), local deviants (four identical tones followed by a different tone; denoted xxxxY), and omissions (four identical tones; denoted xxxx).These stimuli were presented in three types of blocks where one of them was presented with a high frequency (initially 100%, then 75%), while the others were rare. This design thus separated the local deviancy of the fifth sound from the global deviance of the entire sequence, and also allowed to probe whether the omission effect differed when a standard or a deviant tone was expected.

Omission responses might therefore constitute a critical test of the predictive coding framework. Here, we capitalize on omission responses, combined with a hierarchical violation-of-expectation paradigm, to demonstrate that auditory signals are indeed submitted to multiple, hierarchically organized stages of top-down prediction. We use a recently introduced auditory paradigm that can dissociate two types of predictions, based on local probabilities versus global rules (Bekinschtein et al., 2009). In a given block, a frequent sequence of five tones is presented (75% of trials), interspersed with rare violations (15%) in which the frequency of the fifth tone deviates from the expected, and with rare omissions (10%) in which the fifth tone is simply omitted (Figure 4.2-1). Crucially, on some blocks the frequent sequence is of the "xxxxY" type, i.e. four identical tones followed by a distinct one. ERP recordings reveal that the fifth, locally deviant tone, although fully expected, still elicits a MMN. However, only the rare violation sequence, which contains five identical tones "xxxxx", elicits a distinct and later novelty response, the "late positive complex" or P3B wave (Bekinschtein et al., 2009). In the context of hierarchical predictive coding models, this observation can be interpreted as reflecting a "violation of a violation": the monotonous "xxxxx" sequence is surprising because it fails to contain the fifth deviant tone, which normally generates a MMN. Hierarchical predictive coding models thus hypothesize two levels of predictions in this situation: a first low-level expectation, based on local transition probabilities, incorrectly predicts a fifth "x" tone after the first four "xxxx", thus generating a MMN, while a second, higher-level expectation, based on the knowledge of the overall "xxxxY" rule, cancels the surprise elicited by the first level.

A simple prediction ensues. When we omit the fifth sound, thus presenting an identical series of 4 tones "xxxx", the observed omission response should vary according to the expectation induced by the overall context. On "xxxxY" blocks, where two successive predictions are generated, we should observe a large event-related response to omission, composed of superimposed waves corresponding to the predictions of the "x" tone and of the MMN. On "xxxxx" blocks, however, only the first of these predictions should exist, and therefore the event-related response to omission should be significantly smaller. We tested this hypothesis by recording simultaneous EEG and MEG signals to these stimuli, relative to a low-level control where the omission of the fifth tone was entirely expected.

4.2.3 Methods

4.2.3.1 Subjects

Ten healthy subjects (mean age 26 ± 4.5 years; sex ratio 1) with no known neurological or psychiatric pathology were studied. All subjects gave their written informed consent to participate to this study, which was approved by the local Ethical Committee.

4.2.3.2 Auditory Stimuli

Two tones composed of 3 superimposed sine waves (either 350, 700, and 1400 Hz, tone A; or 500 Hz, 1000 Hz, and 2000 Hz, tone B) were synthesized. They were 50-ms long, with 7-ms rise and fall times. Series of four or five such tones were presented via headphones with an intensity of 70 dB and a 150 ms stimulus onset asynchrony. The series could comprise 5 identical tones (local standard, denoted xxxxx), 4 identical tones and a fifth different one (local deviant, denoted xxxxY), or only 4 identical tones (omission, denoted xxxx). Series were presented in semi-randomized blocks of ~3-minutes duration, separated by silences of variable durations (700-1000 ms), during which one series was designated as frequent and the other as rare (see Figure 4.2-1). Each block started with 25 frequent series of sounds to establish the global regularity (global rule). Of the next 100 occurrences, 75% were the frequent series, 15 % the rare series and 10% the omission series. A separate block contained 125 presentations of the omission sequence (expected omissions). Each participant received a total of 14 blocks of 125 trials each (3 replications of the four rules xxxxY and xxxxx with either x=A and Y=B, or x=B and Y=A, plus two xxxx omission blocks with x=A for one and x=B for the other). All stimuli were presented using E prime v1.1 (Psychology Software Tools Inc.).

4.2.3.3 Simultaneous EEE-MEG recordings

Measurements were carried out with the Elekta Neuromag system (Elekta Neuromag Oy, Helsinki, Finland) NeuroSpin, which comprises 204 planar gradiometers and 102 magnetometers in a helmet-shaped array. The built-in EEG system (64 electrodes) was used to record simultaneously EEG and MEG. An electrode on the tip of the nose was used as EEG reference. ECG and EOG (horizontal and vertical) were simultaneously recorded as auxiliary channels. MEG, EEG and auxiliary channels were low-pass filtered at 330 Hz, high-pass filtered at 0.1 Hz and sampled at 1 KHz. The head position with respect to the sensor array was determined by four head position indicator coils attached to the scalp. The locations of the coils and EEG electrode positions were digitized with respect to three anatomical landmarks (nasion and preauricular points) with a 3D digitizer (Polhemus Isotrak® system). Then, head position with respect to the device origin was acquired before each block. Subjects were asked to keep their

eyes open, to avoid eyes movements by fixating a cross, and were constantly reminded to pay attention to the auditory stimuli. At the end of the recording, a questionnaire assessed which regularities and violation types had been detected. All subjects reported detecting both rare sound series and omissions.

4.2.3.4 Data analysis

Signal Space Separation correction, head movement compensation and bad channels correction were applied using the MaxFilter Software (Elekta Neuromag). Principal Components Analysis (PCA) was used to remove EKG and EOG artifacts. Using Fieldtrip software (http://fieldtrip.fcdonders.nl/), trials were epoched from 200ms before to 1300ms after the onset of the first sound, low-pass filtered at 40Hz, and baseline corrected using the first 200 ms of the epoch. After visual rejection of artefacts, the trials were averaged per condition and per subject. The latitudinal and longitudinal gradiometers were combined by computing the mean square root of the signals at each sensor position.

Cluster-based statistics were performed using Fieldtrip software. Statistics were computed between 50 and 250ms for mismatch and omission effects, between 50 and 700ms for the global effect and between 50 and 500ms after the onset of the omitted sound for late omission effect. The threshold was fixed to $p = 0.05$, corrected for the size of the search space (time and sensors). We only report the most significant clusters for each sensor type.

4.2.3.5 Source reconstruction

Anatomical T1-weighted magnetic resonance images (MRI) were obtained for each participant after the MEG experiment with a 3-T Siemens MRI scanner, with a resolution of 1 x 1 x 1.1 mm. Head position indicators and the digitized head shape were used for the co-registration of the anatomical images with the MEG signals. Grey and white matter were then segmented using BrainVisa / Anatomist software package (http://brainvisa.info/). The scalp and cortical surfaces were reconstructed for each subject using BrainStorm software (http://neuroimage.usc.edu/brainstorm/). Models of the cortex and of the head were used to estimate the current-source density distribution over the cortical surface. The forward model was computed using an overlapping-spheres analytical model. The inverse model was constrained to a minimum-norm solution (weighted minimum-norm current estimate, wMNE). Sources were reconstructed at each time point. For each subject, the sources were then projected to a standard anatomical template (MNI). Contrasts between conditions were normalized using z-score normalisation.

4.2.4 Results

We first examined our data for the presence of a local mismatch response evoked by the deviance of the fifth tone. Cluster analysis, implemented in FieldTrip software, was used to identify clusters of neighboring sensors where a significant difference between local standards and local deviants was seen over several consecutive time points (see Methods). This analysis was performed separately for EEG sensors, MEG magnetometers (MEGm), and MEG gradiometers (MEGg), using the root mean square of the two orthogonal gradiometers available at each position (Figure 4.2-2). The results revealed an early effect of local deviance, peaking at around 120 ms after the onset of the fifth deviant tone, and which reached corrected significance for each sensor type (EEG: range of first significant window 85-150 ms, $p=0.002$; MEGg, 80-170 ms, $p<0.0001$; MEGm, 82-150 ms, $p=0.02$).

Figure 4.2-2 shows the corresponding topography and time course of a relevant sensor for each block type. In both xxxxx and xxxxY blocks, the response to local deviance has the topography of a classical mismatch field, with bilateral responses over the left and right temporal regions. The mismatch response was significant and with the same sign in each block (block xxxxx : EEG, 85-180 ms, $p=0.002$; MEGg, 84-150 ms, $p<0.0001$; MEGm, 76-130 ms, $p<0.0001$; block xxxxY : MEGg, 80-140 ms, $p<0.0001$). This response thus indexes a first local level of auditory novelty detection which is blind to global context. Indeed, in xxxxY blocks, although the deviant "Y" sound could be fully expected, the mismatch response to the final Y tone remained. Nevertheless the MMN amplitude was reduced on xxxxY compared to xxxxx blocks (Figure 4.2-2.C; EEG: 134-190 ms, $p = 0.014$; MEGg, 103-700 ms, $p<0.00001$; MEGm, 95-210 ms, $p=0.006$).

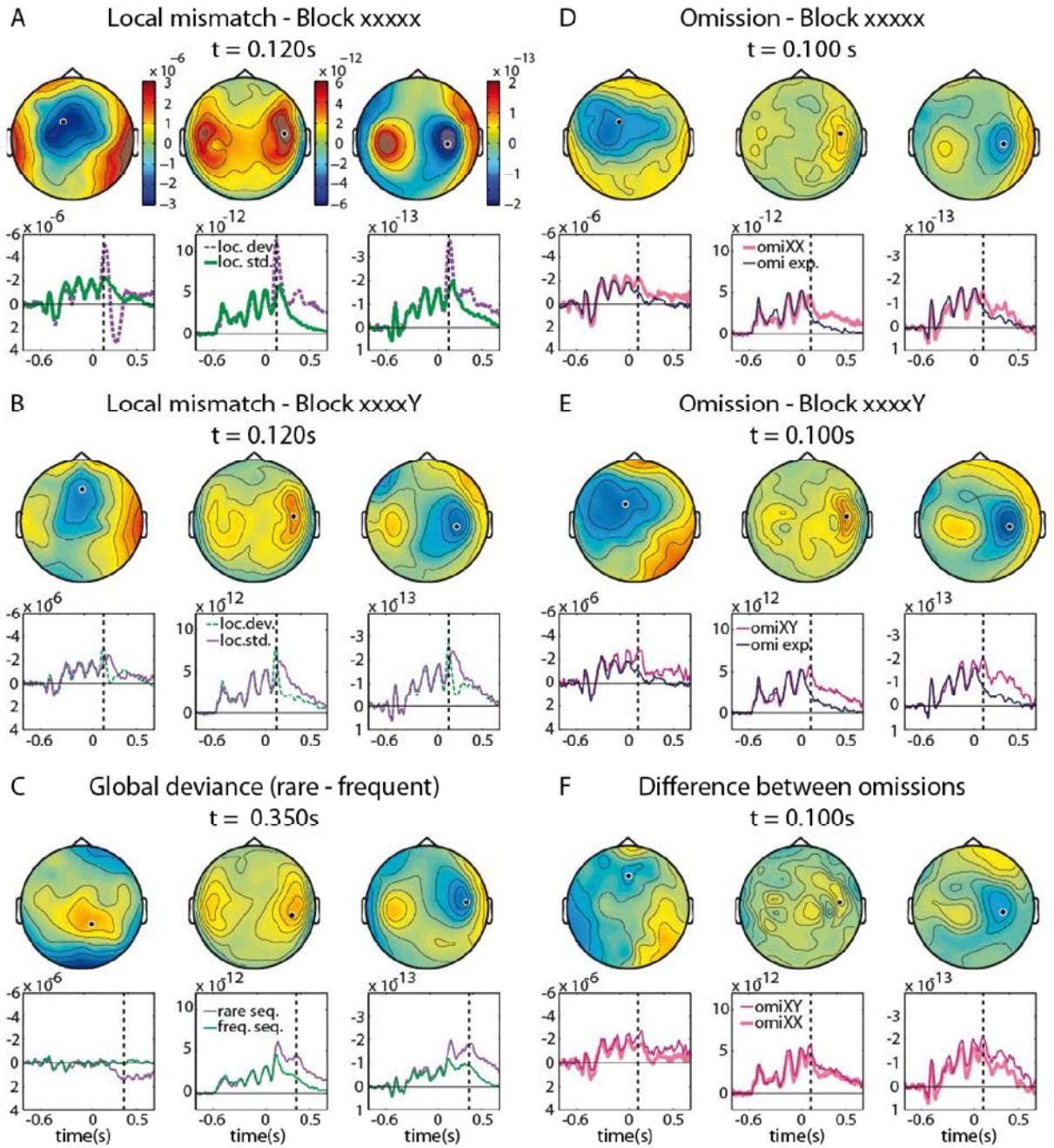


FIGURE 4.2-2: SENSOR-LEVEL TOPOGRAPHY AND TIME COURSE OF THE BRAIN RESPONSES TO DISTINCT FORMS OF NOVELTY.

In each panel, the 3 topographies show the spatial distribution, on a top view of the scalp, of EEG signals (left), MEG gradiometers (norm; middle), and MEG magnetometers (right) at the time indicated (vertical dotted line in the graphs). Graphs show the time course of these signals as recorded from an individual sensor (marked by a dot on the corresponding topographical map). A, B: effects of local mismatch: bilateral auditory areas show a rapid response to the fifth deviant tone, whether it is rare (xxxxx blocks, panel A) or frequent (xxxxY blocks, panel B). C: effect of global deviance: a temporally and spatially extended response, corresponding to the P3b in event-related potentials, is seen in response to rare sequences. D, E, F: responses to omission of the fifth tone (omiXX : rare omissions in the XXXXX block, omi XY : rare omissions in the XXXXY block, omi exp : expected omissions in the control omission block). The brain responds to omission by emitting a sharp response whose amplitude is smaller when a standard was expected (D) than when a deviant was expected (E), resulting in a significant difference (F).

We then examined the presence of a second-level novelty response, dependent on the frequency of the overall sequence rather than of individual tones. On all sensor types, rare sequences differed from frequent sequences on a later time window than the MMN (EEG, 327-540ms, $p<0.00001$; MEGg, 103-600ms, $p<0.00001$; MEGm, 275-600ms, $p<0.00001$). Note that, on xxxxY blocks, this higher-order novelty response was elicited by the monotonic but unexpected xxxxx stimulus relative to the frequent xxxxY stimulus, leading to a complete inversion of the classical mismatch response (Figure 4.2-2B). On such trials, a sequence of two successive novelty events, hereafter termed local and global effects, was thus revealed. In EEG, as previously described, the second, global effect has the classical topography and latency of the P3b component, which differs strongly from the MMN (compare 2A and 2C). Surprisingly, in MEG, these two events have very similar topographies, both dominated by bilateral responses over temporal cortices.

The next step was to examine omission responses. The omission effect was computed by recording the brain responses to rare omissions (presentation of only four identical tones instead of five), separately within xxxxx and xxxxY blocks, and comparing them to a block where only sequences of four identical tones were presented (expected omissions). The results showed an early effect of unexpected omission peaking around 100 ms after the onset of the omitted tone (i.e. 250 ms after the onset of the fourth tone), with a topography similar to the MMN topography for all sensors types (Figure 4.2-2D, 2E). The early latency of this peak response to an absent stimulus is consistent with the hypothesis that this response corresponds to an unfulfilled prediction. The omission effect was significant in both block types (xxxxx blocks: significant only for MEGg, 76-200ms, $p<0.0001$; xxxxY blocks: EEG, 104-160ms, $p = 0.022$; MEGg, 150-200ms, $p<0.0001$; MEGm, 150-200ms, $p = 0.002$). In both blocks, the difference between rare versus expected omissions was also significant in a later time window with a topography similar to the above global effect (xxxxx blocks: EEG, 425-440ms, $p= 0.032$; MEGg, 327-500ms, $p<0.0001$; MEGm, 234-500ms, $p<0.0001$; xxxxY blocks : MEGg, 134-500ms, $p<0.0001$; MEGm, 272-500ms, $p<0.0001$). Thus, the omission effect consists of a sequence of early and late responses, the latter coinciding with the P3B-like global effect observed in all rare conditions (rare sequences and rare omissions). This finding is consistent with the hypothesis that this global effect is a correlate of detection of any deviance from the rule currently entertained in working memory.

Finally, we compared the amplitude of the omission effect between the xxxxx and xxxxY blocks, testing the prediction, unique to hierarchical predictive coding models, that the early

omission effect should be bigger on xxxxY blocks when a deviant stimulus is expected. The difference between omissions is plotted in Figure 4.2-2F. The early omission response was significantly higher in amplitude for xxxxY blocks than for xxxxx blocks (EEG: 109-130 ms, p=0.03, and MEGg, 68-80 ms, p=0.042). Figure 4.2-2D and 2E show the topography of the difference between omissions, indicating a slightly more anterior source for the MEG omission effect on xxxxY than on xxxxx blocks.

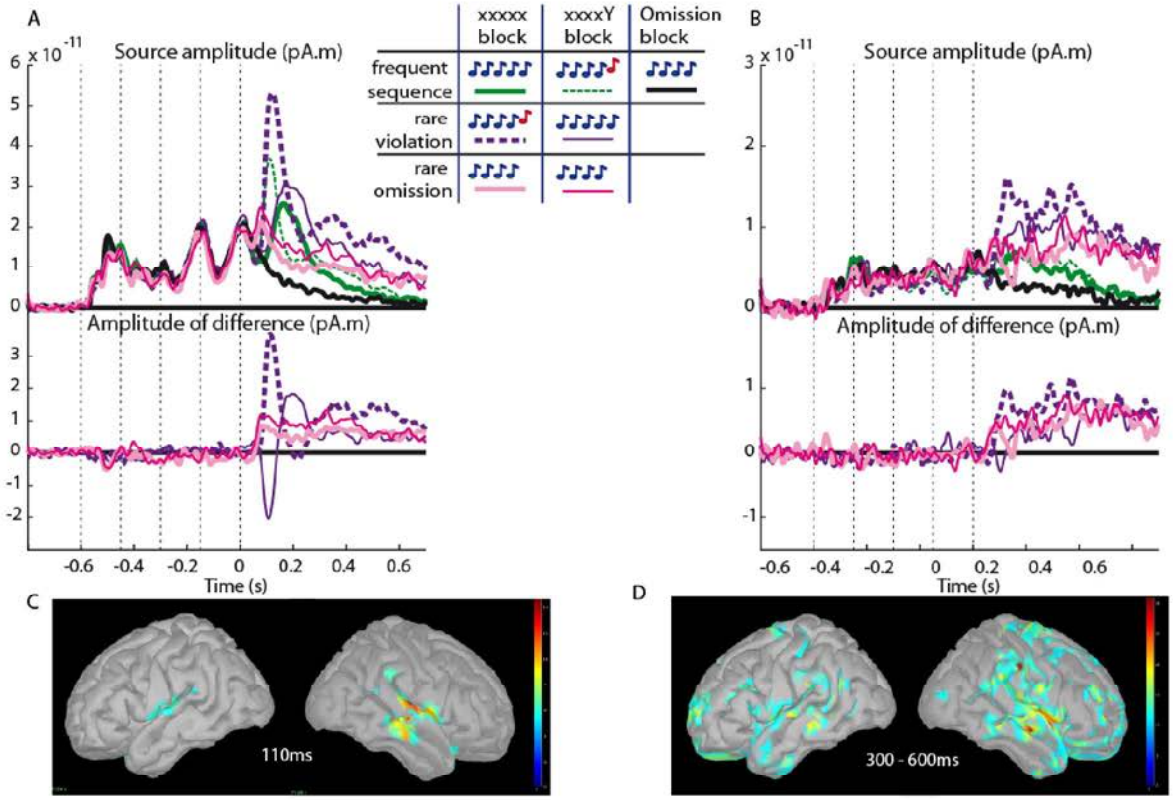


FIGURE 4.2-3: SOURCE MODELING OF THE EFFECTS.

The reconstructed signals from right auditory cortex (A) and right precentral cortex (B) are shown. The top panel shows the signals in each of the seven experimental conditions, and the bottom panel subtractions of each rare sequence (identified by the same color as in the upper plot) minus the frequent sequence of the same block, and of omissions in each block type (same color as in the upper plot) minus the expected omission from the control block. (C and D) z-score corrected source reconstruction of the local effect (C) and effect of deviance from global rule on inflated cortex (D). Local effect (C) is the average effect of deviance from local regularity over both types of blocks at time = 110ms. Global effect (D) is the average of contrasts between rare and frequent trials, at constant stimulus (rare violation - frequent sequences and rare omissions – expected omissions), averaged over the late period of the trials (300-600ms). The source in auditory cortex shows a sharp and rapid response to local deviance, followed by a late and sustained response to global deviance. The source in precental cortex shows only the late sustained response. Both effects are present on omission trials (red/pink curves), with greater initial response to omissions in xxxxY blocks (thin red curve) compared to xxxxx blocks (thick pink curve).

4.2.5 Discussion

By recording event-related potentials (ERP) and magnetic fields (ERF) while manipulating and violating the participants' auditory expectations at two distinct levels, we obtained direct evidence that an active, predictive and hierarchical system underlies the brain's response to auditory stimuli.

We used a minimal norm estimation method to reconstruct distributed cortical sources based on MEG data (similar results were obtained when combining EEG and MEG data). The results dissociated regions sensitive to local and global regularities (Figure 4.2-3). Maximal responses to incoming tones arose from bilateral superior temporal cortices, in the vicinity of Heschl's gyrus, and the underlying segment of the right superior temporal gyrus. These regions also showed the maximal response to local deviants (maximum z-score, in Talairach coordinates: right hemisphere, x = 45 mm, y = -19 mm, z = 13 mm; left, x = -48 mm, y = -16 mm, z = 13 mm) and to omissions (same sources) (see Figure 4.2-3C). Another set of regions did not respond strongly to incoming tones, but responded in a categorical manner to global deviance. The activated sites were highly distributed in bilateral anterior and posterior superior temporal gyri, supramarginal gyri, dorsolateral, inferior, polar and ventromesial prefrontal cortices, anterior cingulate, and the superior parts of the precentral and postcentral gyri (see Figure 4.2-3D). As shown in Figure 4.2-3B, their activity was minimal for frequent sequences, but converged towards a higher and temporally sustained level of activity whenever a rare sequence or omission was presented. Finally, we examined the cortical origins of the difference between the omission effects on xxxxY versus xxxxx blocks. Consistently with the trend seen on sensor-level topographies, the maximal difference between omissions originated from a more anterior temporal region than either the MMN or the basic omission effect, lateralized to the right hemisphere (x = 53 mm, y = -2 mm, z = 4 mm).

First, we replicated the earlier finding of a double dissociation between the early mismatch response (MMN) and a later, temporally extended and distributed response (P3B) (Bekinschtein et al., 2009). The MMN was sensitive to local violations of transition probabilities, and was essentially blind to higher-order regularities, since it continued to be evoked, at a reduced level, by a fifth deviant tone that could be expected (in the xxxxY blocks). Contrariwise, we observed a late (~300 ms) divergence which reflected solely the deviance of the overall sequence rather than of its individual component tones.

While MEG and EEG revealed functionally and temporally similar responses, their spatial pattern diverged. The EEG topographies differed strongly for the MMN and P3 stages,

but in MEG these two stages showed similar topographies involving mainly temporal sources. This difference in sensitivity to sources between MEG and EEG stresses the interest of combining the two methods (Cohen & Cuffin, 1983). Overall, the results suggest an initial stage confined to temporal cortex, and a later stage where this activity is amplified and expands into distributed additional regions, particularly in prefrontal and parietal cortices (Bekinschtein et al., 2009). The weak influence of the latter sources on MEG topography might be due to their multiplicity and dispersion.

It is important to note that the previous results by Bekinschtein et al. (Bekinschtein et al., 2009) in a similar paradigm were observed in the context of a counting task where participants counted the rare stimuli. Thus, the P3b response that they observed on rare compared to frequent trials could have arisen from the counting process, which occurred on global deviant but not on frequent trials. By contrast, the present findings were obtained while the participants' only instruction was to attend to the stimuli. Thus, our results show that the counting task is not necessary, and that the late P3B response reflects, at least in part, the response of a higher-order novelty-sensitive system.

Our findings refine earlier results by showing that the local and global effects are not fully independent (Bekinschtein et al., 2009), but interact in an early time window. Specifically, the local mismatch response was significantly smaller in xxxxY blocks than in xxxxx blocks. There are at least two interpretations of this effect. First, it could be solely due to a difference in transition probabilities. Indeed, MMN amplitude decreases when the probability of the deviant increases, and in the blocs where the xxxxY sequence is frequent, the transition probability x→Y is necessarily higher relative to xxxxx blocs. However, this effect is also fully consistent with the hierarchical predictive coding hypothesis, which predicts that on xxxxY blocks, a second-level prediction can be used to partially cancel out the first-order error novelty response to the expected deviant sound Y.

While the theoretical implications of this early modulation of the initial mismatch response are therefore ambiguous, the complete inversion of the mismatch signals observed in a later time window argues strongly for a hierarchical process. Indeed, on xxxxY blocks, the xxxxx stimulus becomes the rare stimulus and elicits a P3B-like brain response. The fact that a stimulus that consists solely of a repetition of 5 identical tones (xxxxx) can elicit a novelty signal, if the participants expected a different sequence, is in itself highly suggestive that the brain operates as a multi-level predictive system sensitive to prediction errors.

Having established the existence of a hierarchy of at least two novelty systems, we used sound omissions to provide a stronger test of the hypothesis that these novelty responses arise from active prediction systems as opposed to passive neural adaptation (Bendixen et al., 2009a; Friston, 2005; Garrido et al., 2007; Garrido, Kilner, Kiebel, et al., 2009; Winkler, 2007). Our results confirm earlier findings that the omission of an expected tone leads to a time-locked brain response which is easily detectable by MEG and EEG, and has a similar topography as the original evoked response (Hughes et al., 2001; Raij et al., 1997; Yabe et al., 1997). Furthermore, our design tested the novel prediction, unique to the hierarchical predictive coding framework that the omission response should vary with the context. Specifically, this framework supposes that evoked responses reflect a series of prediction errors indexing the difference between the incoming signal and its prediction at successive hierarchical levels. Accordingly, when the incoming signal is omitted, brain responses should reflect solely the predictive signals and how they vary depending on the current context (Bendixen et al., 2009a). In agreement with this notion, we observed that the brain response to an omitted signal, following a strictly identical series of four tones, varied depending on whether the participants expected the fifth tone to be identical or different from the preceding ones (xxxxx versus xxxxY blocks). A significant larger omission response was observed on xxxxY blocks. This difference between the two omissions effects is exactly as predicted by the hierarchical view: on xxxxY blocks, an additional higher-order predictive signal is needed to cancel out the predictable MMN inevitably arising from the novelty of the fifth tone.

Passive adaptation models of mismatch responses attempt to account for omission responses in terms of an oscillatory or rebound response, due to an entrainment of brain oscillators by the rhythm of the preceding stimuli (May & Tiitinen, 2009). This hypothesis, however, cannot explain our observation of a larger omission response on xxxxY than on xxxxx blocks. Under the adaptation interpretation, we would have expected either a constant entrainment by the four preceding tones, and hence a constant omission response; or, if anything, a larger entrainment on the regular xxxxx blocks than on the xxxxY blocks, where the fifth item interrupts the rhythm of the first four ones – exactly the contrary of what was observed. Our results are therefore very difficult to explain with the adaptation hypothesis alone. They do not rule out that sensory adaptation may exist, but only prove that it cannot be the only mechanism at work, as also argued by others (Garrido et al., 2008). Note also that the size of the omission effect goes in the opposite direction as that of the MMN: as described above, the MMN is larger on xxxxx than on xxxxY blocks, but the omission effect is larger on xxxxY than on xxxxx blocks. This inverse relation between the magnitudes of the MMN and of the omission response is

exactly as expected from a hierarchy of predictive systems, but cannot be easily accommodated by a single process of novelty detection. In particular, it rules out the possibility that the observed modulations are due to one of the blocks being intrinsically more interesting, motivating, or attention-grabbing.

The latencies of the observed novelty responses are also indicative of a predictive system. First, note that the timing of the omission response arises too early to correspond to a rebound of a putative oscillation induced by preceding stimuli. As shown in Figure 4.2-3 (left panel, pink curves), the cortical response on omitted xxxx trials does not consist in a series of 5 equally spaced peaks, as would be predicted by the oscillatory adaptation/rebound model. Rather, omission responses arise earlier than the MMN, which itself arises earlier than the response to an expected tone. This temporal order is the opposite of what would be expected from an ascending feedforward system, where the stimulus first has to be processed bottom-up before its departure from the familiar can be detected. It is, however, in full agreement with a hierarchical predictive system where first the presence, then the precise identity of the incoming tones, are successively predicted in advance of the actual stimulus.

The fact that the omission effect is equally early on xxxxx and xxxxY blocks may seem counterintuitive: according to a hierarchical model, one might have expected a sequence of two successive omissions effects. In reality however, although two predictions are indeed assumed, both have to come quite early if they are to act as predictors that cancel out the effects of the incoming signals. Predictive coding models thus predict that, during the xxxxx block, the omission effect must arise simultaneous with the earliest activation evoked by the fifth stimulus. Furthermore, during the xxxxY blocks, an additional second-order omission effect must arise prior to or simultaneously with the MMN in order to act as a predictor of it. The timing of the observed effects is compatible with these hypotheses. Furthermore, their topography suggests that the second-order omission is generated at a distinct cortical site about 2 cm more anterior in temporal cortex.

In summary, in agreement with recent theoretical models of cortical architecture (Friston, 2005; Kiebel et al., 2008; Kiebel, Kriegstein, et al., 2009; R. P. Rao & Ballard, 1999), our findings suggest a hierarchical organization consisting of several successive prediction and novelty-detection systems. The present paradigm, combined with MEG, EEG or intracranial recordings, dissociates at least two levels of prediction: the MMN responds to local auditory predictions while the later P3b responds to more global and integrative violations of expectations. In that respect, our observations add to a growing number of dissociations of these two systems.

Bekinschtein et al (Bekinschtein et al., 2009) demonstrated that the early MMN resists to visual distraction, to non-consciousness of the rule linking the five successive tones, and remains present in coma and vegetative state (Faugeras et al., 2011) while none of these properties hold for the global P3B, which therefore seems to index a conscious process. Prior ERP and fMRI evidence confirms that the superior temporal region can respond to novel stimuli that are subliminal and fail to be detected (Allen, Kraus, & Bradlow, 2000; Diekhof, Biedermann, Ruebsamen, & Gruber, 2009), while a much broader fronto-parietal network, indexed by the P3b, underlies conscious detection (Del Cul et al., 2007; Diekhof et al., 2009; Sergent et al., 2005). Pegado et al. (Pegado et al., 2010) observed that, when the delay between tones is prolonged up to several seconds, the MMN is drastically reduced while the P3b remains constant in size, though slightly delayed, in correspondence with the participants' preserved capacity to detect the violations. Ritter et al. (Ritter, Sussman, Deacon, Cowan, & Vaughan, 1999b), like us, found that the MMN remains while the P3b vanishes in a context where the local auditory deviance is fully predictable (in their case, because it is systematically preceded by a visual cue presented 600 ms earlier).

We conclude that auditory novelty detection appears to be organized in several stages (Winkler, Takegata, & Sussman, 2005). The mismatch negativity reflects the operation of a temporally and conceptually-limited prediction system which uses the recent past in order to predict the present, solely based on a compilation of the probabilities of the stimuli and their transitions. The auditory prediction underlying the MMN may rely on several recent stimuli (J Horváth et al., 2001), but it uses only a limited time window (Pegado et al., 2010; Sabri & Campbell, 2001) and is blind to the global overall "rule" or pattern followed by the stimuli (Bekinschtein et al., 2009). The extraction of such rules and the detection of their violations involve a later, more distributed predictive system (Bekinschtein et al., 2009; Diekhof et al., 2009; Ritter et al., 1999b; Winkler et al., 2005), that sends predictions to the first one in an early time window. The operation of both of these systems is frequently undetectable, as their sole effect is to reduce or cancel the responses evoked by predictable sensory stimuli. The omission paradigm, by unveiling them, provides a flexible method to dissect the brain's multiple top-down expectation systems.

4.2.6 Acknowlegments

This project was supported by a senior grant of the European Research Council to S.D. (NeuroConsc program) and by the Fondation pour la Recherche Médicale 'Equipe FRM 2010' grant to Lionel Naccache. The NeuroSpin MEG facility was sponsored by grants from INSERM,

CEA, the Fondation pour la Recherche Médicale, the Bettencourt-Schueller foundation, and the Région île-de-France, and is part of a general research program on functional neuroimaging of the human brain (Denis Le Bihan). We are grateful to Ghislaine Dehaene-Lambertz, Lucie Hertz-Pannier, Caroline Huron, Antoinette Jobert, Andreas Kleinschmidt and the NeuroSpin infrastructure groups for their help in subject recruitment and testing; Marco Buiatti, Lucie Charles, Sebastien Marti and François Tadel for help in data analysis; and Jean-Pierre Changeux, Ghislaine Dehaene-Lambertz, Alain Destexhe, and Karim Benchenane for useful discussions.

CHAPTER 5

MODELING ACCESS TO WORKING MEMORY AS A SELF-EVALUATION AND DECISION PROCESS

Submitted paper

5.1 Introduction to the article

The two previous chapters I studied the properties of automatic processing of temporal regularities. I showed that the properties of the mismatch response were consistent with a predictive coding model using a short-lasting memory trace.

In this last chapter of my contributions I focus on a different level of temporal regularity processing. In the first chapter of this thesis I showed that conscious access in thought to be characterized by the ignition of a central workspace with strong recurrent long range connectivity. This recurrent connectivity induces a stabilization of the information represented in this global workspace over time and has been described as a close correlate of working memory. Moreover, we showed that learning of long distance dependencies was strongly affected by concurrent working memory tasks suggesting that this correlate of consciousness is crucial form some types of temporal regularity learning.

I also showed in the introductive chapter that working memory is characterized by a limited capacity and a selective access. Understanding how this selective access can be managed to learn temporal regularities in the absence of an explicit task is a necessary first step to study other aspects of conscious processing of temporal regularities.

5.2 Article

5.2.1 Abstract

Working memory offers the unique possibility to maintain information for an arbitrary period of time. However, this ability comes at the cost of a limited information capacity. Hence, the decision to commit a piece of information to working memory or to relinquish it is a crucial one. Here, we propose to model access to working memory as an internal decision process, based on a self-evaluation of the relative values of maintaining or relinquishing information. We further propose that, even in the absence of an external task, the brain manages its memory updating according to some internal goals. We argue that one of these goals is optimization of the prediction of upcoming inputs. We show that using a value system that is sensitive to prediction accuracy, the brain can learn a successful policy to gate access to its working memory system.

5.2.2 Introduction

Although the exact capacity of WM is debated, it is generally agreed that it is very limited. It was first thought to be limited to seven items (G. Miller, 1956) but more recent estimates argue in favor of an even stronger constraint of 3 or 4 items (Edward K Vogel & Machizawa, 2004), organized in discrete slots, that can contain information about multiple features of the same object (Fukuda et al., 2010; Luck & Vogel, 1997). The possibility that it offers to maintain a stimulus for an arbitrary long time is crucial to perform successfully a large number of task, and WM capacity covaries largely with fluid intelligence (Conway, Kane, & Engle, 2003).

The successful use of working memory does not rely solely on the number of items it can hold. Given its reduced capacity, working memory management is needed to filter appropriately the stimuli that are relevant from those that are distractors. (Edward K Vogel et al., 2005) showed that filtering efficiency and working memory capacity were tightly correlated, suggesting that limitations in working memory capacity may in fact reflect the failure to selectively stabilize the relevant information.

A previous model showed that working memory management could be successfully solved using a reinforcement learning approach in the context of complex tasks such as the 1-2-AX task (O'Reilly & Frank, 2006). But even in the absence of explicit rewards, humans still learn about the structure of the world. Responses to unexpected events in a sequence, such as the mismatch negativity (MMN) and the P3 event-related components, indicate that the brain automatically extracts sequence regularities (Huettel et al., 2002). Bekinschtein et al (Bekinschtein

et al., 2009) showed that although the MMN is robust to manipulations of attention, the P3 is strongly affected by the state of consciousness or attention. Moreover, the P3 component is sensitive to higher-level regularities than the mismatch negativity, namely it appeared for violations of expected sequences of sounds contrary to MMN that was elicited by violations of transition statistics between successive tones. These observations suggest that some of the neuronal circuits implicated in the detection of higher order regularities are attention-dependent. A minor variant of the same stimuli sequence (Elyse Sussman et al., 1998) showed that shortening the intervals between sequences makes these regularities accessible to low-level processing units. These results suggest that the key function of attention in this case is bridging the temporal gap between stimuli (Robert E. Clark et al., 2002). Working memory is an obvious candidate to fulfill this function.

In sequence learning paradigms, subjects typically have no explicit task to perform, yet they quickly developed expectations based on the structure of the sequence. We propose that in the absence of an externally guided task, working memory is used to learn how to better anticipate on future events, across arbitrary long temporal gaps. In this paper, we show that an internal self-evaluation system can be defined so that WM can be managed to keep only relevant (i.e. predictive) events in memory and filter out the unpredictive ones, thus optimizing the prediction of future stimuli. The resulting model paints access to working memory as an internal decision process, based on the relative values of maintaining or relinquishing information. We explore the consequences of this management system on the dynamics of memory content and rule discovery.

5.2.3 Methods

5.2.3.1 General notations

The system is submitted to a stream of stimuli. X is the random variable corresponding to the identity of the stimulus.

X can take n values $x_1, x_2 \ldots . x_n$

$X(t)$ is the stimulus at time t.

The system can remember only one past stimulus at a time; One working memory slots contains information about its identity and the time since it was stored, denoted as τ.

$S = \{X(t-\tau), \tau\}$ is the random variable describing working memory state.

A stimulus that is not stored is forgotten forever. After each stimulus, the system must thus choose between two actions: either the current stimulus is stored and the WM is updated ("update" action), or the stimulus previously in working memory is kept and the current stimulus is forgotten ("keep" action).

5.2.3.2 Model description

The system can use the working memory content to evaluate the matrix of conditional probability

$Pe(X|S)$.

This estimate is updated according to the rule:

If $X(t) = xi$

$$Pe(xi|S) \leftarrow Pe(xi\ |S) + \eta$$
$$Pe(X|S) \leftarrow Pe(X|S)/(1+\eta)$$

On average this updating rule produces a convergence of the estimate towards the real probability $Pr(X|S)$ according to the formula:

$$Pe(X|S,n) = Po(X|S) * \vartheta^n + \Pr(X|S) * \eta * \sum_{i=1}^{n} \vartheta^i$$

Where $\vartheta = \frac{1}{1+\eta}$ and n is the number of time the probability was estimated (i.e. the number of time the WM was in state S). Po is the initial estimate of the probability taken by default to be a flat prior and n the number of time the probability was updated.

The system also estimates the average probability $Pe(X)$ using the same updating rule as the $Pe(X|S)$ estimate. This probability is updated at every time step.

The **computational goal** of the system is to maximize the predictive power of the working memory slot, measured by

$$R(t) = log\left(\frac{Pe(X(t)|S(t))}{Pe(X(t))}\right)$$

The main idea is estimate the Value of each state S defined by the average Reward that can be expected under the policy P.

$$V(S,P) = \langle \sum_{t'>t} \gamma^{t'-t} R(t') \rangle$$

Where γ is a constant <1 representing a time discounting (it was taken equal to 0.9 in all simulations)

This quantity Ve is evaluated using a TD rule:

$$\delta = R(t) + \gamma\, Ve\big(S(t+1)\big) - Ve\big(S(t)\big)$$
$$Ve\big(S(t)\big)$$

Where $Ve(S(t+1))$ is the weighted sum of the value of the two possible future states:

$$Ve\big(S(t+1)\big) = P(update|\, X(t), S) * Ve(S = \{X(t), 1\})$$
$$+P(keep\, X(t-\tau)|\, X(t), S) * Ve(S = \{X(t-\tau), \tau+1\})$$

The policy is determined using a softmax function on the estimated values of the two possible future states

$$P(update) = \frac{e^{\beta * Ve(S=\{X(t),1\})}}{e^{\beta * Ve(S=\{X(t),1\})} + e^{\beta * Ve(S=\{X(t-\tau),\tau+1\})}}$$

where $1/\beta$ is a positive parameter called the temperature and regulates the exploration/exploitation trade off.

5.2.3.3 Initial conditions

All initial values are equal, set to a small positive value (0.05 in all simulations). The probability estimate is initialized with a flat prior.

5.2.3.4 Performance estimate

In the first test of the vbWMA model, the optimal policy consists in keeping the predictive events in the memory slot, until it is not predictive anymore. If a predictive event is in working memory and the current stimulus is also predictive, the correct strategy consists in discarding the current stimulus.

We attributed a score to the decision made for each predictive stimulus: If a predictive stimulus was correctly stored until it was predictive or correctly rejected (if another predictive stimulus was in working memory), we scored the decision with 1. If the stimulus was stored whereas it should have been discarded but then kept until it was predictive, we scored this

decision with 0.5. If the stimulus was incorrectly discarded, we scored the decision with 0. Each point on the graph represents the average score over 20 decisions.

5.2.4 Results

Working memory provides a unique ability to maintain selected information for an arbitrary long time. Its capacity limitation poses the computational problem of finding the optimal policy to filter its access. We argue that, in the absence of external feedback, access to working memory is governed by an estimation of the value of stimuli in predicting future events. In other words, working memory is used to optimally anticipate on sensory inputs.

In this paper, for clarity's sake, we study a simplified version of this computational problem (Figure 5.2-1).

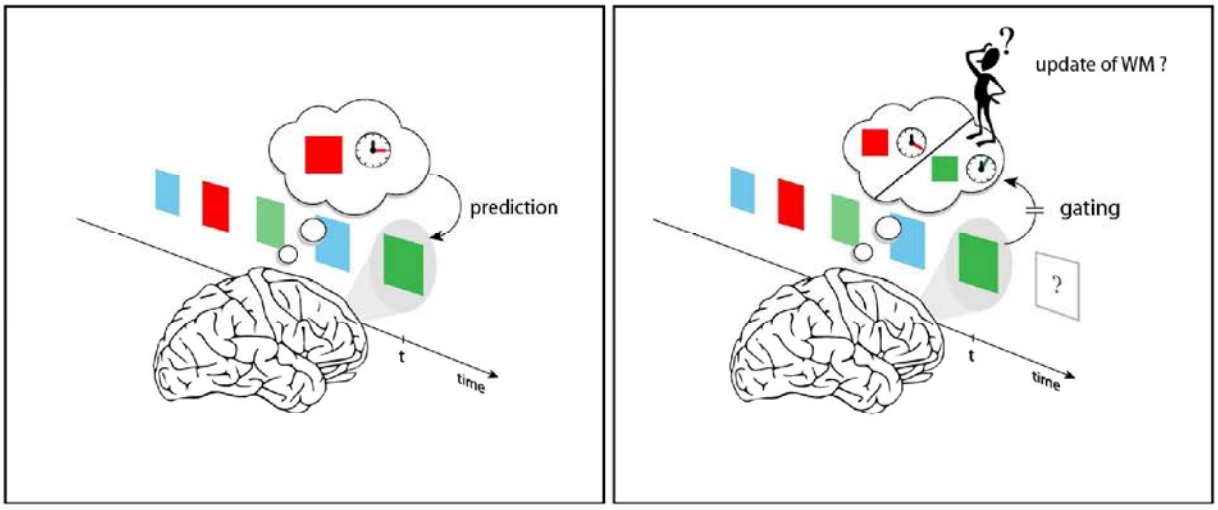


FIGURE 5.2-1: THE SIMPLIFIED COMPUTATIONAL PROBLEM STUDIED.

(left) At each time point, the brain tries and predicts the new stimulus based on what it knows from the past, ie its memory content. Memory content in our case is limited to one of the past stimuli (here the red one) and time since it was presented (here 3 time steps ago). (right) The computational problem is to optimize prediction by deciding for each stimulus whether to store it in working memory or keep the current memory content instead.

5.2.4.1 Simplification hypothesis

We assume that the organism is presented with a time series of stimuli, and tries for each new stimulus to anticipate its identity. This prediction is made according to an estimate of conditional probabilities of the stimuli, given what is known about the past. We assume that the information encoded in working memory is the **identity** $X(t-\tau)$ of the past stimulus that is

maintained and the **time** τ it spent in WM since it was encoded. We assume that time is encoded as the number of past stimuli (independently of their presentation rate).

We reduced the problem to the simple case where working memory capacity is reduced to **one slot** S, with $S(t) = \{X(t-\tau), \tau\}$. This reduces the management problem to a single decision between two choices at each time point: either keep the current memory content or replace it by the current stimulus.

Any item that is not granted access to the memory slot is forgotten. It cannot be used for evaluation of the conditional probability matrix $Pe(X|S)$, nor for prediction of future stimuli.

Note that the working memory capacity limitation is thought to concern only neural representations that are maintained in an active form. Probability distributions and value function can be maintained at no working memory cost as they are hypothesized to be encoded in connectivity, not firing rate.

5.2.4.1.1 General architecture of the model

Framed in this simplified manner, the problem of optimizing working memory usage is similar to an action selection problem: each time a new input is presented; the system must decide whether it should keep the current memory content or discard it to store the current input instead.

This is a complicated problem, because it is necessary to optimize at the same time the exploration of the structure of the world (here the matrix of conditional probabilities) and the exploitation of identified structures to predict next events as best as possible. Even when knowing perfectly the sequence of stimuli, finding the best policy (i.e. the best decision for each couple of current input and current working memory content) is almost always intractable by brute force because of the combinatorial explosion of the number of policies to consider when the number of stimulus identities or the past history considered increases.

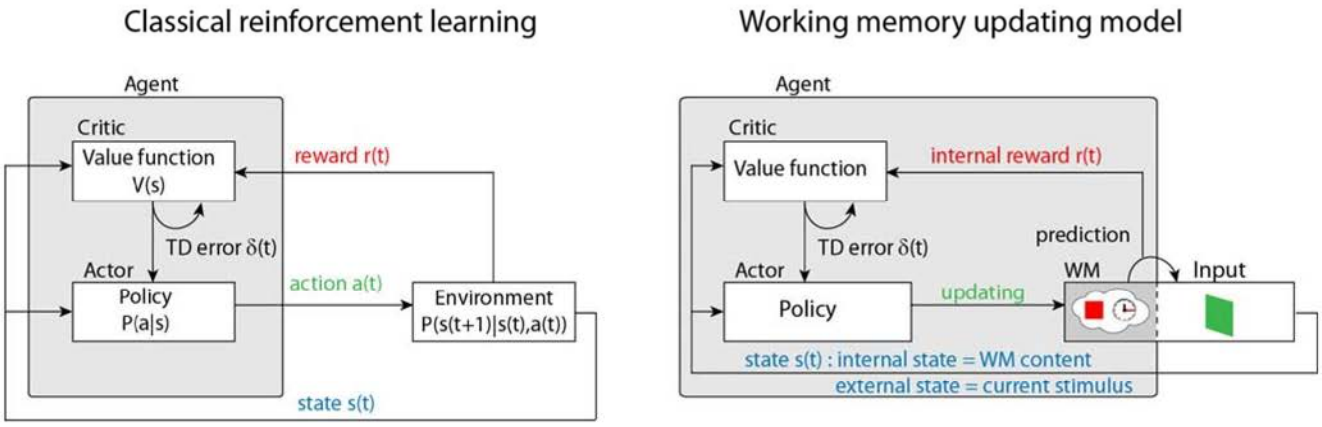


FIGURE 5.2-2: COMPARISON BETWEEN CLASSICAL REINFORCEMENT LEARNING (left) and our model (right). Instead of coming from the environment, the reward is generated internally, and based upon the accuracy with which the input can be predicted. It is used in the same way as in classical reinforcement learning to compute a value function associated with each state, that tracks the reward that can be expected. This value function is used to set an action policy (i.e. a procedure to select an action in response to any possible state). In our case, the policy concerns the internal decision to update working memory content or to leave it unchanged.

Given the similarity of the situations, we propose to solve this problem using a reinforcement learning approach (

Figure 5.2-2). Reinforcement learning is an algorithm that is able to learn from experience which actions lead to desired outcomes. Instead of classically placing the reward in the environment, we only propose that the reward is generated internally according to an internal goal: the prediction of upcoming events. The system is rewarded when the working memory content allow a better anticipation of the next stimulus that a system without such a memory.

Hence, we define the reward at each time as :

$$R(t) = log\left(\frac{Pe(X(t)|S(t))}{Pe(X(t))}\right) \quad (1)$$

Which quantifies to what extent the probability estimate of the current input given what the system has in its memory slot S(t) is superior to the probability estimate without taking into account any history.

A key assumption of reinforcement learning is that past outcome history is a good predictor of future rewards. A value function estimates the average reward that followed each possible state of working memory. In practice it means that a value is associated with each

working memory state (which is defined by the pair {stimulus identity, time it spent in working memory}), and that this value reflects the average reward obtained while having this particular memory content, but also the time discarded average of the rewards that were obtained in the states that followed it.

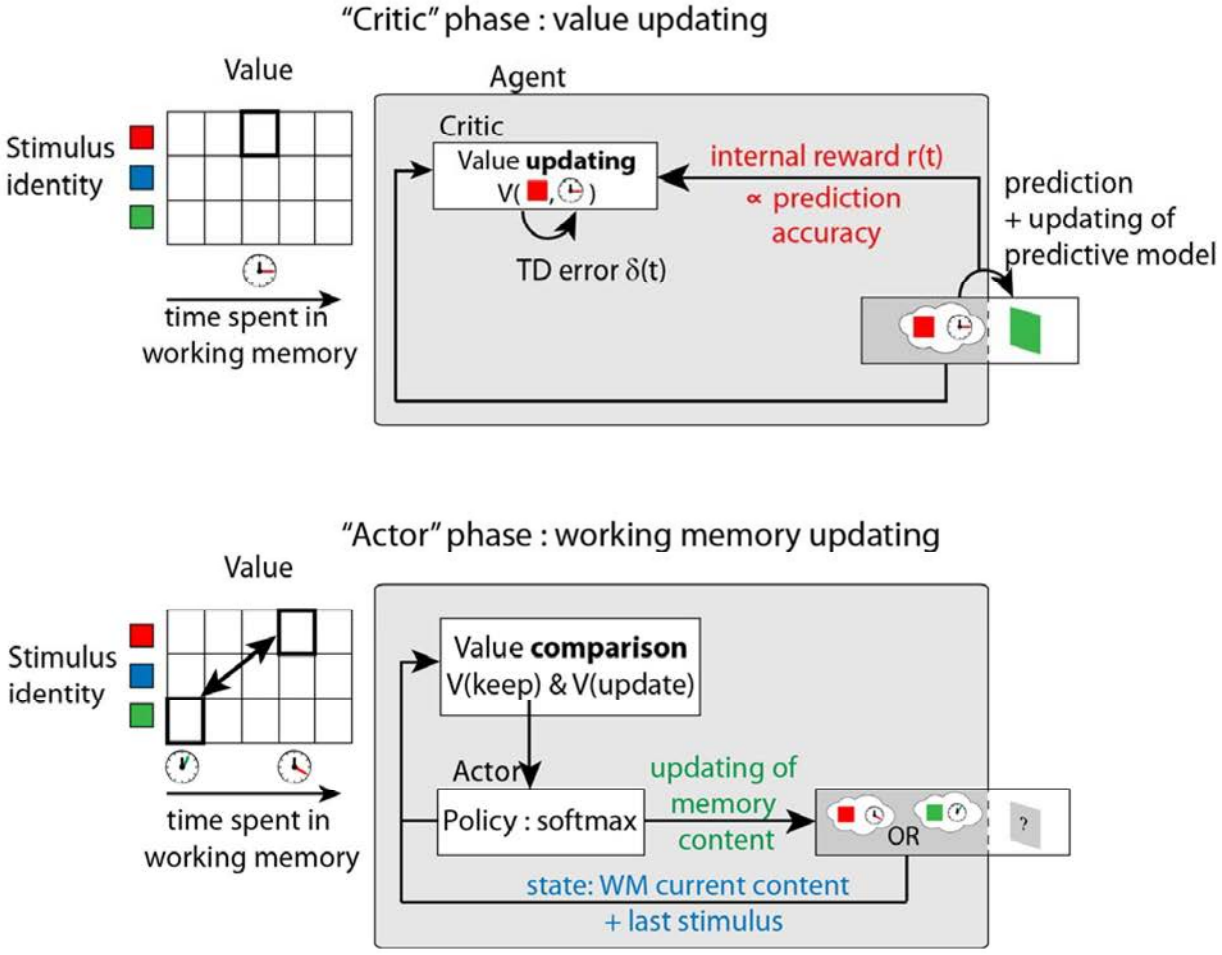


FIGURE 5.2-3: FUNCTIONING OF THE MODEL AT EACH TIME STEP.

(Top) The current memory content and transition probability estimate are used to predict the stimulus. The reward associated with prediction accuracy is computed and used to update the value associated with the current memory state. The conditional probability estimate (probability of each stimulus given the current memory content) is updated. (bottom) the choice between the two possible actions (keeping memory content or updating it with the current stimulus) is made by comparing the values associated with the two states, and choosing the action leading to the state with highest value, with a probability computed by a softmax function.

5.2.4.2

5.2.4.2.1 "Critic" phase (Figure 5.2-3A)

At each time point, the model, hereafter called the "value-based working memory access" (vbWMA) model, predicts the next stimulus based on its memory content and the statistics of conditional probability it accumulated. The reward is computed based on this estimate according

to eq (1). The probability estimate given the current memory state $Pe(X(t)|S(t))$ and base probability $Pe\big(X(t)\big)$ are then updated.

The reward is then compared to the value associated to the current memory state and the value is updated according to a classical temporal-difference learning rule (R.S. Sutton & Barto, 1998a; Richard S. Sutton, 1988)

$$Ve\big(S(t)\big) = Ve\big(S(t)\big) + \eta * \delta$$

$$\delta = R(t) + \gamma\, Ve\big(S(t+1)\big) - Ve\big(S(t)\big)$$

Given this formula, the value associated with one state S does not only reflect the average reward obtained when the memory content corresponds to state S, but also an estimate of the time discarded sum of the rewards obtained in the next states.

5.2.4.2.2 "Actor" phase(Figure 5.2-3B)

Using its value function, the model then decides whether to keep the item currently in memory for the next time step or whether to replace it by the current stimulus. Because the value associated with one state reflects the sum of the future rewards that can be expected, the decision about the updating of the memory content is based on the simple comparison of the values associated with the two possible future states. In practice, the probability of updating the working memory slot with the current stimulus is a softmax function of the difference between the values associated with the state {current stimulus, at time dt =1} and with the state {current memory content, at time dt+1}.

5.2.4.3 Ability to discover relevant dependencies

We first tested whether the vbWMA model could isolate predictive stimuli from a stream of irrelevant ones, i.e. assign a high value to the working memory storage of predictive stimuli. We created a sequence of stimuli where only one stimulus (stimulus 2) acted as a predictor of another stimulus (stimulus 1) with higher probability than otherwise, n time steps later. The probability of all other stimuli was independent of history. We varied the temporal distance between the predictive and the predicted stimuli.

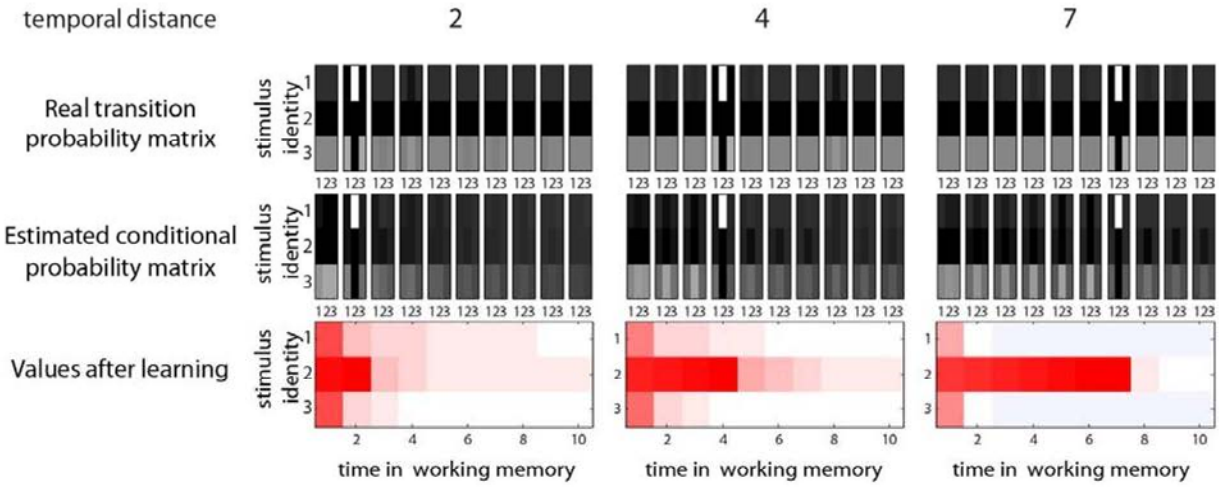


FIGURE 5.2-4THE MODEL DISCOVERS HIDDEN REGULARITIES EVEN AT A LONG TEMPORAL DISTANCE.

(top line) We created sequences of stimuli where the probability of occurrence of each stimulus is independent of the past, except for one predictive relation. In our simulations, we varied the temporal distance of the predictive relation, from two to seven time steps. The top line represents the actual conditional probability of stimuli given past stimulus. (middle line) After presentation of 80000 stimuli to the model, the conditional probability estimate tracks correctly the real conditional probabilities. Notice that probability estimates are better for probabilities conditional upon the predictive stimulus, and for time inferior to the temporal distance of the predictive relation. (lower line) The value function reflects the identity of the predictive stimulus, and the temporal distance of the predictive relation. The transition probability matrix is three dimensional (3x3x10 for this figure). Each subpanel shows the (3x3) transition probability given what happened t time steps before.

The probability distribution estimate at the end of the simulation shows that the model explored the conditional probability matrix so that the predictive relation was correctly estimated. As a result, a high value is assigned to the states where the predictive stimulus is in working memory are high. We can see (Figure 5.2-4) that the states that have high value are those that precede the predictive time, and that this value function accurately follows the true conditional probability matrix, for temporal distances between the predictive stimulus and predicted time up to seven time steps. We can also observe that the conditional probability estimates are closer to the real conditional probabilities when they depend on the predictive stimulus. This is due to the fact that the predictive stimulus was more often in working memory, and as a result was estimated more often.

5.2.4.4 Ability to develop the right strategy

We then studied the dynamics of policy discovery. In this simple case, the optimal strategy to exploit the structure of the stimuli sequence for better prediction is straightforward: the predictive stimulus should be kept in working memory until the time where it is no longer

predictive. In the case where both the stimulus and the memory content are predictive, the memory content should not be replaced.

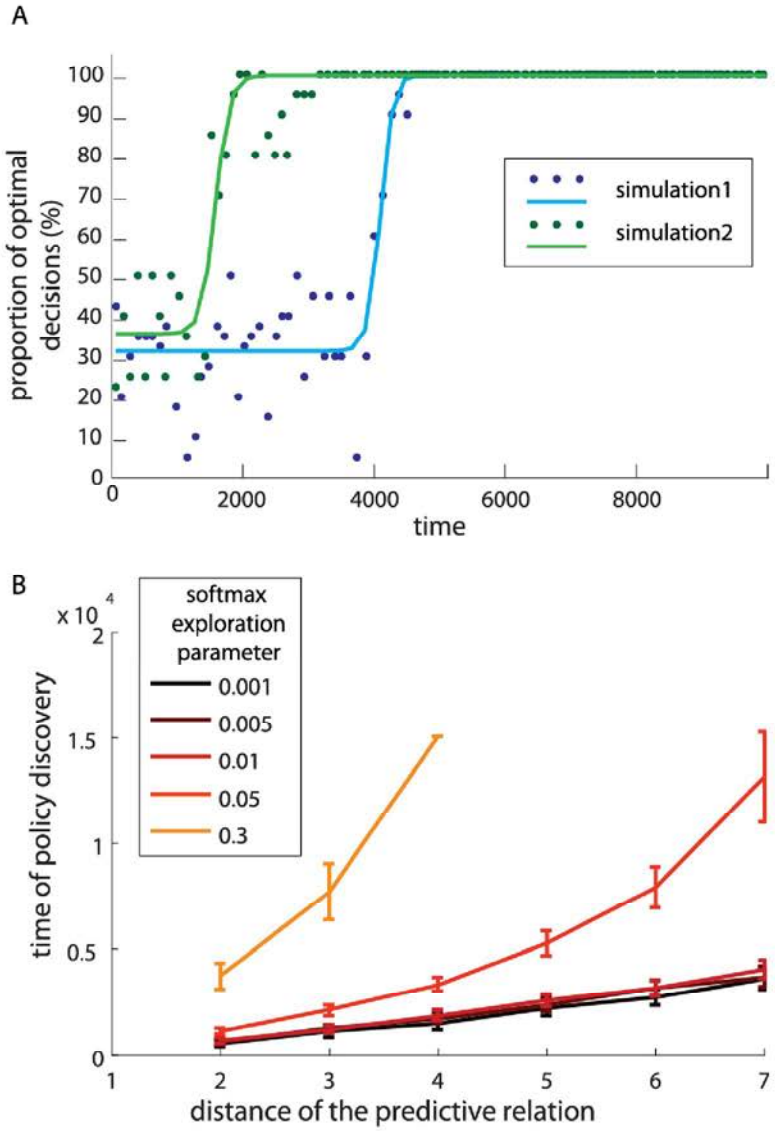


FIGURE 5.2-5: DYNAMICS OF DISCOVERY OF A SUCCESSFUL POLICY.

(A) A successful policy is discovered in a non linear way. Example of two simulations on sequences of stimuli generated according to the same rules. The behavior can be modeled as a sigmoid function, and the inflection point is defined as the time of policy discovery. (B) The time necessary to discover a successful policy increases with the temporal distance of the predictive relation. It also increases when the exploration parameter beta of the softmax function increases. Each point represents the mean and standard deviation of the time of discovery for 20 simulations.

Figure 5.2-5A shows the fraction of optimal decisions made regarding the predictive stimulus for two simulations. The dynamic of the discovery is nonlinear and can be modeled by a sigmoid function: there is a long period without much learning, followed by a relatively sudden discovery of the predictive item and therefore of the appropriate policy. This non-linearity is due

to the positive interaction between probability estimation and value: once the probability estimate for the predictive stimulus becomes slightly informative, its value increases, which increases the probability of keeping the interesting stimulus in memory, making the probability estimate more accurate and increasing the reward. Once the strategy is discovered, it remains optimal for the rest of the simulation.

5.2.4.4.1 Variability of the speed of discovery

We then studied how the time of discovery varies with the delay of the predictive relation, for different values of beta, the parameter of exploration from the softmax function. The higher beta, the higher the probability of choosing the option with the lowest value. In this simple case, the discovery of the appropriate strategy is higher for smaller beta (Figure 5.2-5B). Because there is only one interesting event, the more exploitative the model is, the faster it will converge towards the solution. Empirically, we observed that the discovery time T increases with the delay of the predictive relation (dt), and follows tightly (R>0.98) a relation

$$T(dt) = \ C * \left(\frac{1}{p}\right)^{dt}$$

where p is a constant inferior to 1. The constant C does not vary monotonically with the exploration parameter. p decreased when the exploration parameter increased. To get a intuition about this relation, we can consider that the time of discovery is tightly related to the time necessary to get a good estimate of the conditional probability given the predictive stimulus at the delay dt. It takes an approximately fixed number of evaluations of the probability estimate to make it informative enough to trigger the strategy discovery. Before the strategy is discovered the probability of keeping the relevant stimulus is about 0.5 at each time point (updating or keeping the current content are equally likely). Therefore, the probability of keeping the stimulus for dt time steps is proportional to $(0.5)^{dt}$. The average recurrence time of this state of working memory is $\left(\frac{1}{0.5}\right)^{dt}$ which is why the time to rule discovery follows a rule of that form. In practice, p is higher than 0.5 because the probability of keeping the relevant stimulus starts to increase slightly before the point we called rule discovery.

5.2.4.5 Dependency between speed of rule discovery and divergence between the probability distributions

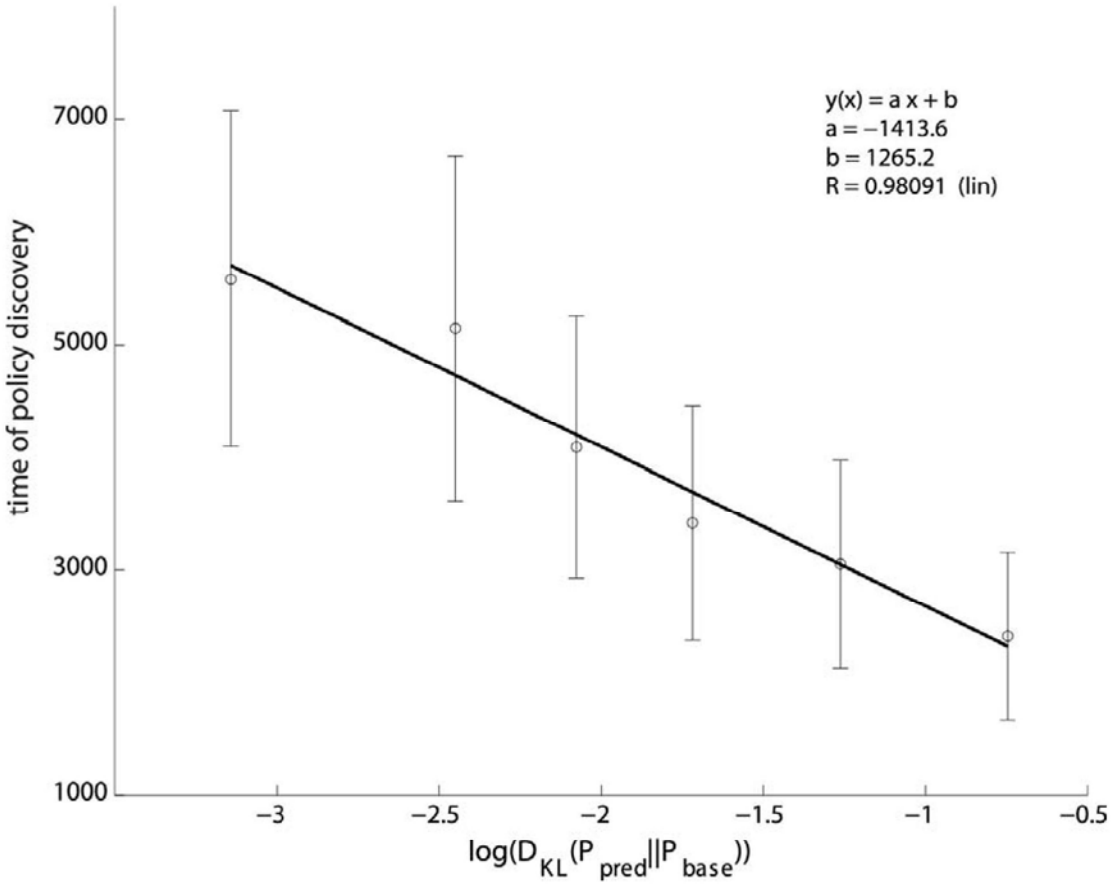


FIGURE 5.2-6: A LAWFUL RELATIONSHIP RELATES THE TIME OF SUCCESSFUL POLICY DISCOVERY TO THE INFORMATIVENESS OF THE PREDICTIVE STIMULUS.

Informativeness is measured as the Kullback-Leibler divergence (D_{KL}) between the unconditional probability distribution of the stimuli, and their conditional probability distribution given the predictive stimulus. We performed 600 simulations with randomly generated probability distributions and sorted them by their degree of informativeness. Each point shows the average (and standard deviation) of the time of policy discovery for 100 simulations whose median informativeness is plotted on the x axis.

An equivalent way of saying that stimulus 2 predicts that stimulus 1 will occur with a higher probability is to say that stimulus 2 predicts a change in the probability at the delay of the predictive relation. We varied randomly the distance between the probability distribution given the predictive stimulus and the base probability distribution, for a fixed delay of 4 time steps between the predictive stimulus and its predictive time. We observed (Figure 5.2-6) that the time of discovery of the strategy was negatively correlated with the log of the divergence between the predicted probability and the base probability (Kullback-Leibler divergence D_{KL}). The more similar the two probability distributions are, the longer it takes to discover the strategy. We also

observed that when the probability distributions are more similar, the vbWMA model fails more often at discovering the strategy.

5.2.4.6 Comparison with other models

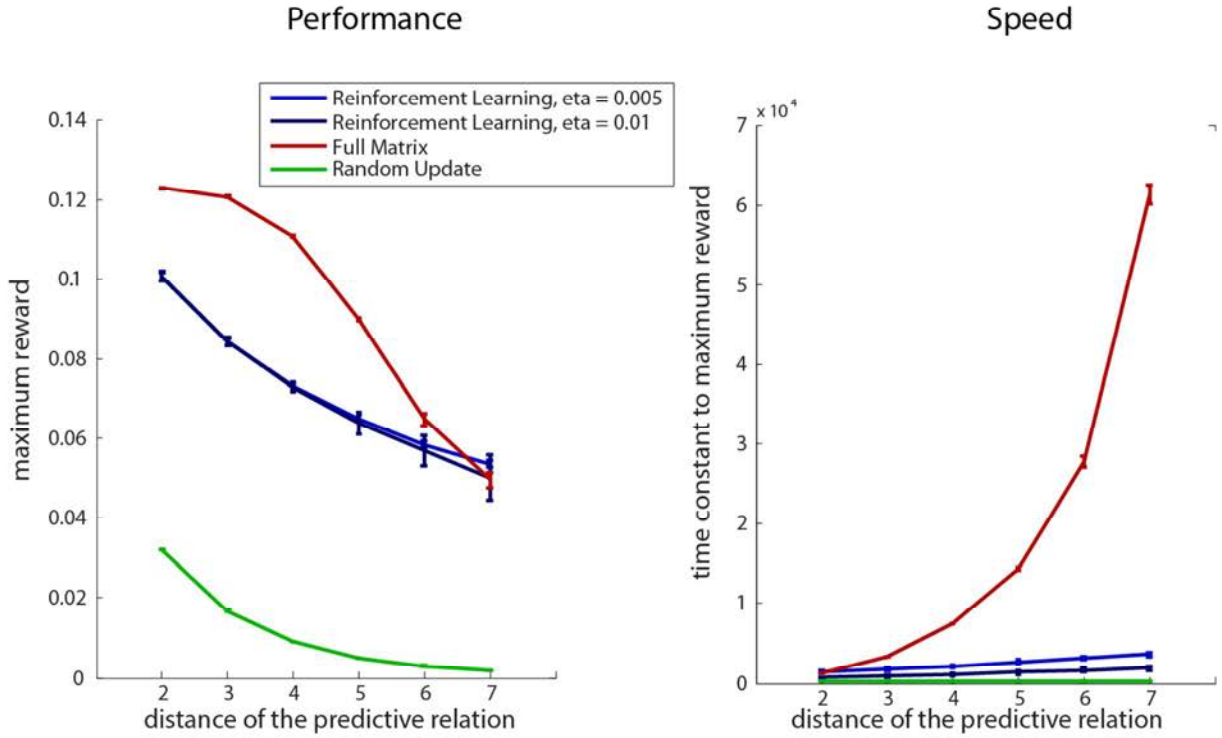


FIGURE 5.2-7: THE PRESENT MODEL ACHIEVES A GOOD COMPROMISE BETWEEN FINAL PERFORMANCE AND SPEED OF LEARNING, ESPECIALLY FOR LONG DISTANCE PREDICTIVE RELATIONS.

Four models are compared: the reinforcement learning presented here, with two learning rates (eta), a model that keeps the full relevant history ("full matrix"), ie the last h items if the distance of the predictive relation is h, and a model that has the same capacity as the reinforcement learning model but updates it randomly. In order to estimate the average reward after learning and the learning speed we fitted the reward collected as a function of time with a saturating exponential. (Left) Performance is considerably improved compared to random update, especially when the temporal distance of the predictive relation becomes large, and is only slightly worse than the full matrix that represents an upper bound on the performances. (right) Final performance is reached up to an order of magnitude faster using reinforcement learning than using the full history.

In order to demonstrate the usefulness of our model's components, we compared the performance of the vbWMA model to two other models with a simpler architecture. First, a "full matrix" model keeps the entire relevant history, i.e. if the predictive relation is at 4 time steps, it keeps the last 4 stimuli in memory, and estimates the full conditional probability matrix $P(X(t)|X(t-1), X(t-2), X(t-3), X(t-4))$. This model gives an upper bound on the reward that can be collected, as it eventually tracks the actual generative model of the stimuli. However, its needs in

memory capacity become very large for long time dependencies; because the conditional probability matrix that it estimates is (n+1) dimensional if n is the delay of the predictive relation.

We evaluated the effect of strategy discovery by comparing the vbWMA model to a "random update" model with similar memory capacity (a single slot) but updating the memory at random (while equalizing the average time spent in WM). Figure 5.2-7 shows the amount of reward collected per time unit once the models reach an asymptotic regime. Our model performs much better than the "random update" model, showing the importance of having a strategy for working memory management. Our model does not quite reach the performance of the "full matrix" model, as it is more constrained. However, when looking at the time required in order to learn the predictive structure and reach the maximum reward, we can see that our model learns much faster than the "full matrix" model to get to its optimum. Indeed, when the predictive delay becomes large, the full matrix becomes exceedingly large (of size n^{dt+1}), making the estimate of all the possible states very slow. In summary, the vbWMA model realizes an excellent compromise between final performance and learning time.

5.2.4.7 Learning of more complex sequential regularities

We then looked at the vbWMA model behavior for more complex types of regularities. We generated sequences where 10 stimuli were overall equally likely, but comprised several distinct regularities: (A) stimulus 2 predicted stimulus 1 after 3 time steps; (B) stimulus 2 predicted stimulus 3 after 5 time steps; and (C) Stimulus 4 predicted stimulus 5 after 3 time steps. Although Stimulus 2 predicted both stimulus 1 and stimulus 3, stimulus 1 was not predictive of stimulus 3.

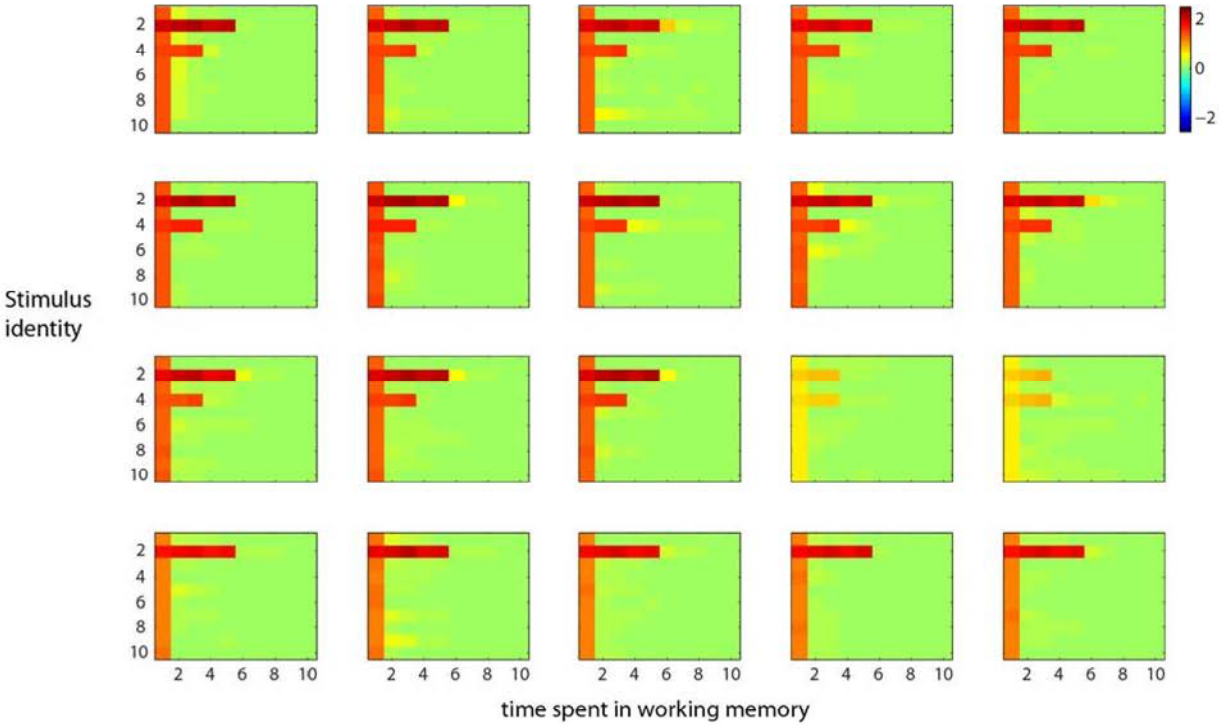


FIGURE 5.2-8: VARIABILITY IN THE POLICY DISCOVERED ON 20 DIFFERENT RUNS OF THE SAME MODEL.

We run 20 simulation of the model on sequences generated according to the same rules: stimulus 2 was predictive in two predictive relations of temporal distance 3 and 5, and stimulus 4 was predictive in a relation of temporal distance 3. The results were reordered to regroup similar results. In 13 simulations, all the predictive relations where discovered. In 2 simulations, only the short distance relations were discovered, and in 5 simulations, only the predictive relations involving stimulus 2 were identified.

Figure 5.2-8 shows the values at the end of 20 simulations (beta =0.01). We can see that the model did not behave the same way for all simulations: the discovery of predictive relationships was all-or-none and stochastic. In 65% of the simulations, the values reflect the discovery of all the predictive relations. Notice that the values associated with stimulus 2 before 5 time steps are all larger than the values associated with stimulus 4. It means that a stimulus 4 in working memory will be maintained unless a stimulus 2 occurs, which is the appropriate strategy. In 10% of simulations, only regularities A and C, which involve a 3-time-step predictive relationship, was discovered. The exploitatory policy limited the exploration of the later time steps strongly enough so the 5 time steps predictive relation was never discovered. In the remaining simulations, only the predictive relations relying on stimulus 2 were discovered, and the policy that developed favored stimuli 2 to occupy almost always the memory slot, preventing sufficient exploration of the conditional matrix for the other stimuli. These two failures to discover some of the predictive relations reflect the exploration/exploitation tradeoff problem of reinforcement learning. The number of failure can be reduced by increasing the exploration

parameter (we observed only one failure with beta =0.05), but at the cost of a longer average time of strategy discovery.

5.2.4.8 Learning of the Local-global paradigm

One of the motivations for our model was to capture the Bekinschtein et al (Bekinschtein et al., 2009)paradigm. In this study, sequences of five tones where presented to subjects. In a crucial condition, most sequences where composed of four identical tones followed by a different one (AAAAB sequences). Rarely, this rule was broken by the presentation of sequences of five identical tones (AAAAA sequences). Data showed that early auditory areas where unable to learn the regularity, and a mismatch negativity (MMN), marker of deviance detection in these regions, continued to be elicited by the rare B tone. A response to rare sequences of 5 identical tone was observed later in time (around 300ms), corresponding to the P300 potential. Crucially, this response was only observed in conscious and attentive subjects. Similar sequence, with shorter and regular delays between B tones did not elicit a mismatch negativity (Elyse Sussman et al., 1998), suggesting a crucial role of conscious storage of information in the Bekinschtein paradigm to enable the discover the higher order rule.

To test this hypothesis, we trained the vbWMA model on series of sequences containing 80% of AAAAB sequences and 20% of AAAAA sequences, the model developed a strategy where the B stimulus has a high value for 5 time steps, and a positive but lower value for the next 5 time steps. It results in a policy where only B stimuli are stored and they are replaced by the next stimulus B. The fifth stimulus of the sequences is predicted to be a B as the conditional probability matrix given B predicts a B with 0.8 probability at 5 time steps. It would then reproduce the prediction error response observed to the sequence AAAAA in this context. However, the time to converge towards the final strategy is about 200 sequences, which is much longer than observed in the data.

5.2.5 Discussion

5.2.5.1.1 Summary

Working memory has a limited capacity, which imposes to develop a strategy to optimize its usage. In this paper, we explored, on a simplified case, the hypothesis that efficient management of memory content can be achieved autonomously in order to meet the internal goal of anticipating on predictable events. We showed that an algorithm based on reinforcement learning principles, but using internally generated rewards can discover predictive relations at a long temporal distance, and can learn to selectively store predictive stimuli in working memory

for the appropriate duration. We studied the dynamics of strategy discovery and showed that it is a nonlinear process that depends on the delay between predictive and predicted events, and on the Kullback Leibler divergence between the probability distribution.

5.2.5.1.2 Neuronal mechanisms of working memory management

The vbWMA model makes three important predictions regarding the neuronal mechanisms that may underlie optimal working memory management. First, it predicts that the valuation system, involving the striatum and dopamine can be influenced by internal rewards and intrinsic goals. Second, it proposes that correct anticipation of future events is one of these internal goals. Third, we argue that this system plays a determinant role in working memory updating.

These predictions could be put to a test using the capacity of functional magnetic resonance neuroimaging to estimate the degree of activation of valuation networks in the brain, particularly in the ventral striatum where a quantitative relationship is found between activation and subjective reward (Schultz, 2010).We predict that this system will be active, not only in externally rewarded task, but also during self-motivated spontaneous behavior such as sequence learning. We predict that predictive events will activate these systems more strongly after learning than before. We also predict, at the behavioral level, that predictive values and external rewards should combine for determining the preference for access to working memory.

Although the appropriate tests of those predictions remain to be performed, the idea that rewards frequently arise from a self-evaluation process rather than from the environment and can be used to achieve internally determined goals is becoming increasingly influential in robotics (Herd, Mingus, & O'Reilly, 2010; Oudeyer, Kaplan, & Hafner, 2007; S. Singh, Barto, & Chentanez, 2004; Satinder Singh, Lewis, & Barto, 2009) but also in neuroscience (Dayan, 2012; Stanislas Dehaene & Changeux, 1991). In this model we pushed the idea further by applying it to a completely internal process with no action on the external environment. Considerable evidence supports the idea that the basal ganglia/dopamine circuit implements reinforcement learning in the brain (Schultz et al., 1997). Dopamine and basal ganglia have also an important role in working memory maintenance and selectivity, by helping to stabilize persistent activity associated to maintenance of information, and by contributing to the suppression of interference by distractor stimuli (Durstewitz, Kelc, & Güntürkün, 1999; Durstewitz, Seamans, & Sejnowski, 2000; Gao, Krimer, & Goldman-Rakic, 2001; Gruber, Dayan, Gutkin, & Solla, 2006b; McNab & Klingberg, 2008; Müller, von Cramon, & Pollmann, 1998; van Schouwenburg, den Ouden, &

Cools, 2010). These data make plausible the idea of management of working memory content through this neuronal circuit.

We also argue that predicting future events is one of these internal goals, and that the ability to predict induces a self-generated reward. Although we are not aware of data acquired in the absence of external reward, some experimental data argue for an intrinsically rewarding value of information(Behrens, Woolrich, Walton, & Rushworth, 2007). In a task where the expected reward is the same, monkeys preferred the situation where the cues about the amount of reward are the most informative (Bromberg-Martin & Hikosaka, 2009) , which is not predicted by simple reinforcement learning algorithms. Another line of evidence comes from the phenomenon of ambiguity aversion (Hayden, 2010) where less predictable situations seem to be less desirable, even when they have a greater expected objective value. This paradox can be resolved if the predictability of the outcome is seen as a reward in itself, that combines with the external reward.

5.2.5.1.3 Partial implementation of curiosity

The hypothesis that organisms may be intrinsically motivated to search for regularities in the environment bears an evident relation with the notion of "curiosity", and indeed our model provides a partial implementation of this concept. However, in the present model, curiosity is not permanent, but is limited to the initial stages of learning. In our network, we initiated the value matrix not with a zero prior, but with small positive values. This can be interpreted as a positive bias towards unexplored states: states that have not been explored are considered interesting until proven otherwise. After learning, a limitation of the vbWMA model is that it will prefer completely predictable situations to situations where something new can be learned. Overcoming this limitation would require to add an automation mechanism similar to the one described for habit formation (A. Graybiel, 2008) where the dopamine system would disengages from the action selection process after stabilization of the policy.

5.2.5.1.4 The importance of being small

The constraint on memory capacity that is observed in biological systems lies at the foundation of the optimization problem that we studied. The limited capacity of working memory has been seen both as a weakness and as a strength (Nelson Cowan, 2010). Indeed, comparing the vbWMA model with the "full matrix" model revealed an interesting effect: having a limited memory size limits the size of the explored state space, thus reducing the combinatorial explosion. Our simulations prove that, when supplemented with an appropriate access management system, a limited working memory size can provide a good compromise between

final performance and speed of learning. This property, that could be called the "importance of staying small" (Elman, 1993) might be an important feature of working memory and could explain why it is advantageous not to have a capacity larger than a few slots.

5.2.5.1.5 Sudden rule discovery

We observed a discovery dynamic that is a mixture of a slow linear step, the estimation of the conditional probability matrix, and a nonlinear step (Dayan, 2007), the decision to update or not working memory. As a result, we observed that the behavior of our system was non-linear, with a sudden jump in performance. This nonlinearity in discovery is consistent with the dynamics of learning curves in biological processes that often show abrupt transitions between behaviors rather than a graded convergence towards a final state (Gallistel, Fairhurst, & Balsam, 2004). This nonlinearity is partly responsible for the speed up in reward collection, as it favors the exploration of the interesting states as soon as the smallest evidence of predictive power is uncovered. The probabilistic decision process that generates the nonlinearity also introduces stochasticity in the discovery process. As a result, repeated simulations on identical sequences may result in different strategies. In particular, when multiple relations are present in the input stream, the vbWMA model may become blind to some of them, because the exploitation of the first regularity prevents sufficient exploration of other options. Although suboptimal, this behavior might actually reflect a real constraint on learning behavior.

5.2.5.1.6 Engagement as a default behavior

We chose to consider only two possible actions at each time step: storing the current stimulus in working memory, or relinquishing it, leaving the current content unchanged. One could argue that maintenance in working memory has a cost in itself and that a third decision should be added: disengage working memory altogether. We chose not to consider this option, first because it would only worsen the performance of the model, second, because neuronal data (Meyer, Qi, & Constantinidis, 2007) reveal an increased firing rate during the inter-stimulus delay in lateral prefrontal neurons coding for the previous stimulus, even in the absence of an explicit delayed-response task in monkeys, and even in the absence of predictive relations. These data support the model's hypothesis that working memory is constantly engaged, be it only to store the last item seen.

5.2.5.1.7 Relations with conscious access

Working memory is thought to be one of the essential properties of the "global workspace" that underlies conscious processing (B. J. Baars & Franklin, 2003; S. Dehaene &

Naccache, 2001; Stanislas Dehaene & Changeux, 2011; Stanislas Dehaene et al., 1998b). In this respect, the present model may be considered as a partial implementation of the gating function that controls access to consciousness. Conscious access has been viewed as an internal decision process (S. Dehaene, 2008; Shadlen, 2011), namely the "decision to engage" the full resources of the organism, and particularly the working memory resources that rely primarily on dorsolateral prefrontal cortex and associated areas. Here we propose specifically that this function is implemented by a non-conscious valuation system which chooses to bring an item to the conscious foreground as a function of an internal evaluation of its relevance. The relation between dopamine and consciousness is supported by data (Van Opstal et al., 2014) showing a direct correlation between individual variations in striatal dopamine and conscious access to visual information. Crucially, note that according to our model, the decision to access one item rather than another is *not* itself conscious, but only the result of that decision, i.e. access of the memorized items to working memory, is. The hypothesis of a modulation of access to working memory by a value system is also consistent with the phenomenon of attentional blink (Raymond, Shapiro, & Arnell, 1992), whereby having a relevant item in working memory temporarily blocks access to consciousness for other subsequent items. Experimental evidence that, even during this period, the incoming stimuli remain unconsciously evaluated (Anderson, 2005) and their target status is noted (Marti, Sigman, & Dehaene, 2012).

5.2.5.1.8 Limits and future directions

This work is only a first step in the understanding of autonomous management of memory content. We used several simplification hypotheses that make quantitative prediction of the model hard to test in human subjects. First, we considered only one slot of information, where the current estimate of working memory capacity is about 3-4 objects (Nelson Cowan, 2010; Edward K Vogel & Machizawa, 2004). Having more slots opens new possibilities, like the discovery of predictive relations depending on the conjunction of two events, but it brings also a number of new computational challenges that are beyond the scope of this paper. Second, we do not take into account other properties of working memory like chunking or compression of information into more effective representations. Indeed, other groups have tackled the information processing limit not by solving the selection problem, but by trying to understand how to re-encode the stimuli, thus producing a lossy compression of the initial information (Tishby, Pereira, & Bialek, 2000). These two kinds of processes are most likely complementary rather than opposite approaches.

One of the limits of the present model is that, although it is able to learn the global sequence "AAAAB", it would be completely unable to discover that this trial has the same structure as CCCCD, where A,B,C,D are distinct sounds. Rule-governed generalization is impossible in the current framework because we do not consider the possibility of representing more abstract information about the stimuli than their identity. In particular, there is strong evidence that relational properties between stimuli are represented, for example, their same-different status or similarity with respect to the current content of working memory (Engel & Wang, 2011; E. K. Miller et al., 1996). The possibilities offered by this new kind of representation to learn elementary abstract rules (Marcus et al., 1999) will be the subject of future work.

Note that in the present work, knowledge of predictive rules was assumed to be encoded implicitly in synaptic weights. This choice has two consequences: all possible rules are evaluated in parallel and at no working memory cost, but at the same time, only one set of rules can be learned. Some evidence exist that abstract rules are actually represented in the firing rate of a population of prefrontal cortex neurons (Shima, Isoda, Mushiake, & Tanji, 2006). In the future, considering this possibility might make rule acquisition and application more flexible.

Acknowledgment

We thank Rava Da Silveira, Wieland Brendel, Michael Berry and Adrienne Fairhall for usefull discussions.

CHAPTER 6

GENERAL DISCUSSION

In this thesis, my goal was to better understand the processing of temporal regularities, and in particular the differences between the type of computations and neuronal architectures involved in conscious and unconscious processing of these regularities. I used combined modeling and neuroimaging techniques to investigate the type of computations underlying automatic and attentive temporal regularity processing. In this discussion, I will first summarize the main results of this thesis. Then I will discuss the interplay between computational principals that may be shared across all temporal regularity learning processes, and principals that could be specific to conscious processing. Finally, I will consider the limitations of my work within these questions and open some computational perspectives to resolve them.

6.1 Summary of the main results

I identified the mismatch response (MMN) as a representative response characteristic of the unconscious processing of temporal regularities. Indeed, the presence of a mismatch response is not affected by attention, or top down predictions which makes it a pre-attentive component characteristic of an automatic process. Moreover, while it is mainly studied in the auditory modality, counterparts of the MMN have been described in other modalities, suggesting that the computations that give rise to the MMN response might be of larger interest than purely auditory processing. I developed a spiking neuron model based on predictive coding principles and showed that it could reproduce the main properties of the mismatch response. Specifically, I proposed that some neurons in supragranular layers of the auditory cortex present a predicitive activity that can be used to cancel out predictable inputs that arrive to layer 4, so that only prediction error is encoded at this level. I proposed that the "internal model" used to drive the activity of the predictive populations is implemented by synaptic weights between the predictive populations and dynamic attractors that form a memory trace of the past stimuli for a few hundreds of milliseconds. I showed that the prediction error signal could be used as a "teacher" signal to drive NMDA-dependent plasticity between the predictive populations and the memory trace to learn the appropriate internal model of the temporal succession of stimuli. I showed that

an asymmetric STDP rule similar to the one described by (G.-Q. Bi & Poo, 1998) reinforced maximally synapses between the predictive populations and the neurons from the memory trace that were firing on average right *before* the predictable stimuli. After learning the activity from the predictive population has therefore the appropriate timing to cancel out the predictable inputs right before they arrive. I showed that the synaptic weights that constitute the internal model follow closely the conditional transition statistics between the inputs. The following table summarizes the MMN properties that were accounted for by my model and the main properties of the computational network that is responsible for it.

MMN property	Model property
MMN elicited by rare stimuli in the oddball paradigm	The internal model captures conditional transition statistics
MMN increases when proportion of deviant decreases	
MMN can be elicited by repetition in alternate sequence	
Occurrence of a rare tone at regular positions in a sequence ceases to elicit MMN if the SOA is short	Conditional transition statistics capture "long distance" dependencies that have a fixed timing
… but NOT if SOA is longer	The memory trace has a limited duration
MMN source comes mainly from supragranular layers	Canonical microcircuit architecture
MMN is NMDA-R dependent	The STDP is NMDA dependent
MMN is sensitive to recent history in random sequences	Online learning by STDP captures local variations in transitions probability
Omission of an expected tone elicits a novelty response	Predictive population
This omission response peaks earlier than the MMN	The asymetry of the STDP rule, driven by prediction errors during learning produce a predictive activity that peaks right *before* the expected stimulus
A change in ISI or duration elicits a MMN	Neurons in auditory cortex have onset/offset responses + the STDP rule induces timing specific predictions

A key prediction of my model is that the MMN is elicited by violations of transition probabilities. The main alternative model of the MMN, the habituation model, states that the

MMN is the result of synaptic habituation, which is sensitive to frequency of occurrence of a sound. I used these two predictions to propose a protocol that would decorrelate maximally probability of occurrence and probability of transition. Specifically, I presented most of the time pairs of two different sounds AB (frequent transition from A to B), very rarely (a pair every 10 to 20 s). The two models made qualitatively opposite predictions. The results were consistent with the predictions of the predictive coding model, ruling out the pure synaptic habituation account of the MMN.

The MMN model predicted that omission response reflect pure predictive responses. We reasoned that if the computations that generate the MMN are not specific to primary auditory cortex, but constitute a more general principle of temporal processing, we should be able to observe additive omission response when multiple levels of temporal regularity processing are necessary to cancel out incoming inputs. We adapted a paradigm using sequences that could be regular at multiple hierarchical levels. We predicted that in blocs where two levels of regularity where necessary to describe the sequence, we would observe a larger omission response. This prediction was validated by the empirical data.

In the last part of my thesis I focused on one property of conscious processing: the access to the possibility to hold a limited quantity of information for an arbitrary long time in working memory. I explored the computational properties of temporal regularity processing using such a memory form. In particular, the combination of potentially indefinite maintenance with highly limited capacity creates a challenging decision making problem to optimize the way this specific form of memory trace can be exploited. I proposed that efficient management of the working memory content can be achieved by using a value system that tracks the predictive power of stimuli in order to select into memory the items that lead to better predictions of the next elements of a sequence. This model relies in particular on the hypothesis that the accuracy of predictions in sequence processing is considered by the brain as an intrinsic goal that generates an internal reward. This model was able to maintain the appropriate items in working memory to exploit long distance dependencies between elements of a sequence distant of as much as seven time steps.

6.2 Discussion

Given the results of this thesis, what are the common computational principles that underlie temporal regularity processing? What can be the specificities of conscious processing?

6.2.1 General principles for temporal regularity learning

The main shared principle across all chapters of this thesis is certainly the implication of an active predictive mechanism in temporal regularity processing. In both models, the computational goal of the network was the prediction of future inputs. The existence of a predictive activity involved at multiple levels of temporal regularity processing was confirmed by neuroimaging data in *chapter4*.

6.2.1.1 Roles of predictions and prediction errors

In the introduction of this thesis, I identified mainly three roles for predictive activity in temporal regularity learning: avoiding the encoding of redundant predictable information, the anticipation and better identification of future events and appropriate decision making in order to optimize future outcomes. All three of these functions had a role in the models presented here. Both models use predictions for the anticipation of the next stimulus. Both models also use prediction errors, which implies that predictions were subtracted from incoming inputs consistently with the efficient coding principle. It is interesting to note that in the model developed in chapter5, prediction are actually used at two levels: by the predictive mechanism in working memory that tries to anticipate the next stimulus based on working memory content and by the valuation system, that tries to evaluate which of the current memory content or the current stimulus will allow the best predictions in the future to make the appropriate decision.

The second model showed clearly how reinforcement learning mechanisms could be implicated in a predictive process. However, even in the first model, the learning rules used to modify predictive synaptic weights were very close from associative learning principles : prediction error were used as a supervisor for learning from a memory trace, just like reward prediction errors are thought to drive plasticity with neurons from an "eligibility" trace. Associative learning and predictive coding may be two ways of describing very similar computations, the first one emphasizing on the supervisory role of prediction error in driving learning and the second one insisting on the fact that prediction error encodes surprise (den Ouden et al., 2009).

6.2.1.2 Learning implicit models of the world

In both models presented in this manuscript, the internal representations of regularities were extremely similar: they relied on conditional transition probability encoded in synaptic weights. The global neuronal workspace hypothesis states that only information that is represented explicitly in neuronal firing can be conscious (S. Dehaene & Naccache, 2001). As a result, this type of internal representation of sequential dependencies would be implicit, both in

the MMN and in the working memory dependent models, which makes the prediction that while the rules may be learned behaviorally, the subjects should not be able to report verbally what they have learned. In other words, they should not have a metacognitive access to the rule they learned: if the stimulus2 predicts that stimulus1 is very likey to occur 4 time steps later, the predictive mechanism I proposed will be able to *learn* the rule in the sense that it will be able to correctly anticipate the occurrence of stimulus 1, but there will be no unit coding explicitly for the rule: "stimulus2 predicts stimulus1 in four time steps". This prediction is consistent with the fact that long distance dependencies can be learned implicitly but that this learning is affected by concurrent demanding tasks (Curran & Keele, 1993; Remillard & Clark, 2001; Remillard, 2010).

Learning predictive models using implicit representations such as synaptic weights is advantageous because it allows a massively parallel evaluation of hypotheses. Bayesian ideal observer models typically define a hypothesis space, compute the likelihood of each hypothesis given all previous data and then used the Bayes rule to evaluate the likelihood of the data given the posterior distribution over hypotheses. In the MMN model, each synaptic connection between a neuronal assembly encoding a past stimulus and the predictive populations can be seen as the estimated likelihood of the hypothesis "stimulus X, n time units ago predicts stimulus Y". I showed how these weights can be adjusted online and in parallel. This mechanism is therefore an interesting candidate for a biologically plausible implementation of Bayesian inference. Notice that a remarkable property of this implementation, that could be tested experimentally, is that it predicts that the more recent history has more weight in the computation of the posterior distribution over hypotheses. Given the success of optimal observer models to explain behavior (Bejjanki, Beck, Lu, & Pouget, 2011; M. C. Frank & Tenenbaum, 2011; Mars et al., 2008; Orbán, Fiser, Aslin, & Lengyel, 2008; Pouget, Deneve, & Duhamel, 2002; Téglás et al., 2011), it is possible that this type of computation is a general principle for the learning of temporal predictions, with different types of hypotheses spaces tested in different regions of the brain, in function of the types of representations that are encoded at in each area.

6.2.2 Specificity of conscious processing

The initial motivation of this work was to better understand which type of temporal regularity processing would produce neuronal responses that could be considered reliable signatures of consciousness. In this section I will review the main distinctions between conscious and unconscious processing observed in my models. I will also explore how the format of representation of stimuli that have been described in working memory could allow new kinds of

temporal regularity processing. I will also discuss for each of these properties the type of computational challenges they imply and possible options to resolve them.

6.2.2.1 Exact timing and event based timing

In this thesis, I proposed that the possibility to maintain information across time was an important aspect of conscious processing. One of the predictions of the model is that working memory enables the learning of conditional transition statistics in a manner that is much less dependent of the particular timing than in early sensory cortices. The sensitivity of predictions to exact timing is consistent with the existence of a MMN to changes in ISI and was also recently demonstrated in visual cortex (Gavornik & Bear, 2014). The event-based encoding of time that can be observed in working memory (E. K. Miller et al., 1996) does allow more time invariant representation of sequential transitions of probability. The prediction that conscious processing allow generalization of a rule to new temporal intervals, while unconscious processing allows the learning of time-interval specific rules is currently under experimental investigation.

Prefrontal cortex neurons were also shown to encode sequential order in working memory (Mushiake, Saito, Sakamoto, Itoyama, & Tanji, 2006; Ninokura, Mushiake, & Tanji, 2003, 2004; Sawamura, Shima, & Tanji, 2002). This type of encoding would allow prediction based on the sequential ordinal position rather than prediction based on transitions (Endress & Wood, 2011; Orlov, Yakovlev, Hochstein, & Zohary, 2000).

6.2.2.2 Abstract rules

In the previous sections, we showed that the implicit representation of rules or hypothesis was an efficient implementation strategy for parallel estimation of the likelihood of a large number of hypotheses. This encoding strategy presents however the important drawback that only one set of rules can be encoded in such a way. Neuronal data also show that sequential rules can be encoded explicitly in prefrontal cortex in an abstract way regarding both identity and time (Shima et al., 2006). Specifically, Shima and collaborators trained monkey to perform a task where multiple sequences of four movements were possible (Figure 6.2-1). The sequences could be constituted of two repetitions of an action, followed by two repetition of another action (AABB), or present an alternation of two actions (ABAB) or four repetitions of the same action (AAAA). A large proportion of neurons in the lateral prefrontal cortex were found to encode the abstract repetition structure of the whole sequence during the preparatory delay preceding the behavioral realization of the sequence of actions. This code is also a temporal abstraction as a sequence extended in time and representing multiple items is summarized as one object in the firing of these neurons. I will first consider the implications of the encoding of abstract relations

between the elements of the sequence have on regularity processing computations, then I will consider how temporal abstractions can be used and created.

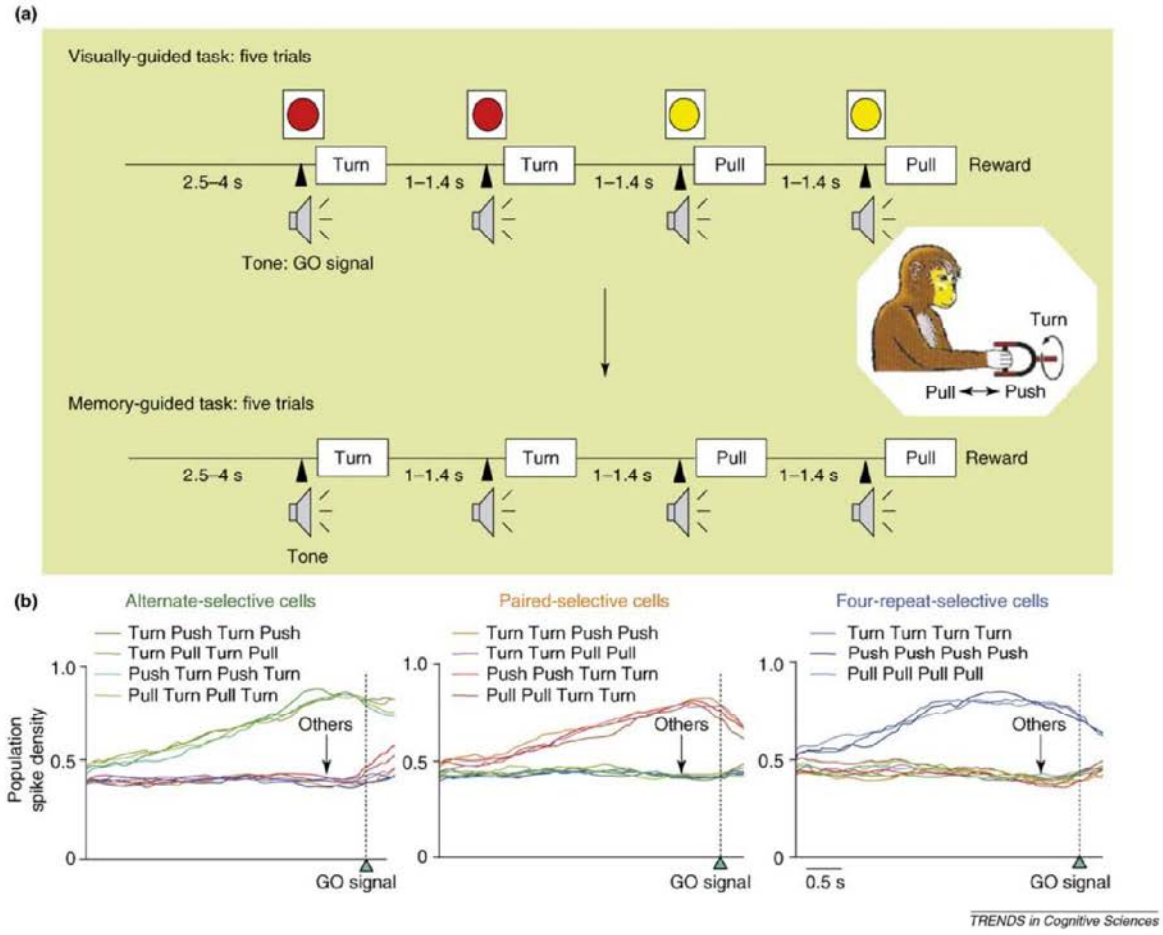


FIGURE 6.2-1 : EXPLICIT REPRESENTATION OF AN ABSTRACT RULE IN LATERAL PREFRONTAL CORTEX (LPFC).

(a) In a motor task, monkeys had to perform a sequence of 4 actions (Turn, Push or Pull a lever) in three possible orders: alternate two actions (ABAB), perform twice the first action and then twice the second (AABB), repeat four times one action (AAAA). (b) spike density of the 22% of the recorded cells in lPFC that showed selectivity to one of the abstract pattern ABAB, AABB or AAAA during the delay period preceding the performance of the task. Note that the activity of the populations does not depend on the particular actions that constituted the pattern. From (Tanji, Shima, & Mushiake, 2007)

6.2.2.2.1 Using abstract relations to predict the next elements

The data from (Shima et al., 2006), show that abstract relations of identity are explicitly encoded in prefrontal cortex. The working memory dependent repetition enhancement observed in prefrontal cortex (E. K. Miller et al., 1996) can be used to learn to represent explicitly same/different relations (Engel & Wang, 2011). In theory, the model we developed in chapter 5 could be extended to predict whether the next object is going to be same or different relative the object stored in working memory, instead of predicting a specific identity. For example, we could

estimate $P(same|\{id, dt, n\})$, the probability that the next object will have the same identity than the one currently stored in working memory, in function of the memory content that tracks the identity id of a past object, the time dt since it was encoded, and the number of time n it was repeated since its encoding. Estimating this conditional probability is not different from what could be done in chapter5, it just uses different representations. However, adding the possibility to encode the same/different relation between successive objects poses two new computational challenges. First, a new dimension is added to the stimulus encoded, and it is clear that some of the dimensions may be irrelevant. For example, using the marginal P(same|{dt, n}) may allow for generalization of a repetition-based pattern to new stimuli. However, it is unclear how the choice of marginal should be done. Second, two types of probability, conditioned on the same dimension of the stored stimulus, can now be used to predict the next stimulus: the probability of a given identity or the probability of a repetition. How can we arbitrate between the two probability estimates?

A possible solution to the second problem might reside in estimating the precision, i.e. the degree of uncertainty about the two estimates and use the most reliable at any given time. This type of solution was suggested in reinforcement learning to arbitrate between the two competing value systems in decision making: model based and model free estimation (Daw et al., 2005). Note that the hypothesis space of the same/different estimate can be much smaller than the precise identity estimate in realistic situations where more than two objects can occur. The P(same) distribution should therefore be reliably estimated much faster than the P(id) distribution. Moreover, when new stimuli are presented, only the P(same) distribution contains some information, and should therefore be relied on in priority in new contexts, allowing the automatic generalization of previously learned repetition patterns to new sets of stimuli.

The coexistence of multiple rules or features creates a new computational problem: what should be the rules to update the value of an object? If the predictive rules do not depend on the identity of the stimulus, it seems unreasonable to update values in an identity specific way. The value learning algorithm should therefore be informed of the predictive rule that were used so that value is updated along the same dimensions. As in the model from chapter5 we predict that if multiple rules involve different stimuli over the same delays, the rules will be in competition with each other and some of them may be missed. Evidence for such phenomenon in abstract rule learning can be found in (Gerken, 2006): infants were familiarized with a continuous stream of syllables following the AAB structure, where A and B are syllables drawn from two pools of syllables (e.g. *lelediwiwijededeje…).* In the test phase, the infants preferred CCD compared to CDC

structures where C and D are syllables drawn from a new pool of syllables. This reproduced the result from (Marcus et al., 1999), showing that infants are sensitive to the abstract repetition structure. Another group of infants was familiarized with AA**b** sequences where A was drawn randomly from a pool of syllable and **b** was a fixed syllable (e.g. *lele**di**wiwi**di**dede**di**…).* In the test phase they were presented with CCD and CDC structures similarly to the previous group. They did not prefer one type of structure over the other, suggesting the adding the rule "**b** is repeated every 3 syllable" blocked the acquisition of the repetition rule. This blocking can be explained if the children adopted the memory managing policy "keep **b** in working memory for 2 syllables then replace it with the next **b**" which led them to ignore any regularity occurring on the intermediate syllables.

6.2.2.2.2 Temporal abstraction: chunking and explicit testing of hypothesis

The fleeting nature of the memory trace used in the MMN model imposes an online treatment of the incoming input, with no time to consider multiple options. On the contrary, the working memory enables maintenance of information so that multiple hypotheses could be considered to infer the correct rule underlying the generation of the sequence of inputs. Note that to achieve efficient coding in the predictive coding framework, one only need to learn over time a good predictive model. If an input was not predicted before the stimulus occurred, it the prediction error should be used *i)* to update the inference about higher order causes so that next stimuli can be predicted correctly and *ii)* update the generative model of input so that occurrence of the same history will not elicit a prediction error next time it is encountered. Crucially, all the computations regard the prediction of future stimuli, not past ones. In some domains like linguistic however, inferring the right underlying structure for every stimulus is essential. A grammatical mismatch response has to lead to a revision of the current hypothesis regarding the underlying syntactic structure to reinterpret the whole sequence of past stimuli. For example, in a garden path sentence such as "The horse raced past the barn fell", the reader usually first parse the beginning of the sentence as a noun phrase plus an active verb. As a consequence, the last word "fell" is unexpected in this structure and should raise a syntactical prediction error signal. However, in linguistics, getting a prediction error, use it to update an internal model and move on to the next stimulus hoping that the model is going to predict better next sentences is not satisfactory. It is necessary to *reconsider* all past stimuli to *test* other syntactical hypotheses and find one that is consistent with the whole sentence.

We showed that the implicit representation of rules or hypothesis was an efficient implementation strategy for parallel estimation of the likelihood of a large number of hypotheses.

This encoding strategy presents however the important drawback that only one set of rules can be encoded in such a way. A potential solution to this problem would be to add "context" units that would create multiple orthogonal codes for a given information in function of the context, so that different rules could be learned regarding the succession of the same stimuli in different contexts. The fact that neuronal selectivity in prefrontal cortex is highly variable in function of the task (Asaad, Rainer, & Miller, 2000; Sigala, Kusunoki, Nimmo-Smith, Gaffan, & Duncan, 2008; Warden & Miller, 2007) is consistent with this hypothesis.

The previous section explored some of the challenges linked to the complexity of the objects represented in working memory that would be necessary to consider while moving towards more realistic accounts of conscious temporal regularity processing. However, this challenges are still compatible with a non-hierarchical and implicit representation of the temporal regularities. Yet, the data from Shima et al. are only one of the evidence showing a hierarchical organization of representations. This hierarchy is typically referred to as "chunking", meaning that an entire sequence of stimuli is represented as one object. Very little is known about the mechanisms that lead to the creation of chunks although they are thought to rely on basal ganglia (A. M. Graybiel, 1998). Interestingly, defining a relevant chunk is formally similar to the problem of creating an option in hierarchical decision making (Dezfouli & Balleine, 2013; Richard S. Sutton, Precup, & Singh, 1999), i.e. creating meaningful temporal abstraction that can then be treated as one object by value systems. This idea of hierarchical decision making using options is the object of active research in hierarchical decision making (Botvinick, Niv, & Barto, 2009). One of the main challenges in this domain is to understand the rules that govern the creation of such chunks or the identification of relevant subgoals in a task. At the neuronal level, the dorsolateral prefrontal cortex (DLPFC) has been shown to be involved in the representation of task sets (Hoshi, Shima, & Tanji, 1998), which are formally similar to options. According to the guided activation theory (E. K. Miller & Cohen, 2001), prefrontal representations do not implement policies directly, but instead represent "context" units that allow the representation of multiple stimulus response mapping. Similarly, chunks could represent context units that allow multiple memory content - prediction associations.

However, the explicit representation of the rules, forbids the use of the parallel mechanism previously suggested to update the posterior distribution over hypothesis. Context units and rules associated with them would have to be tested serially to see how well they explain the sequential data. A similar problem has been described with place cell in the hippocampus. The places cells have a spatial receptive field that depends on the context (the cage) where the

animal evolves, in the same way that prefrontal neurons have stimulus selectivity that depends on the context. When a rat is moved to a new environment, it is necessary to infer the right context so that the position encoding can be expressed by place cells. During the brief period where the animal is uncertain about the correct contextual hypothesis, place cells have been showed to oscillate between the codes for the two possible contexts following a theta rhythm (Jezek, Henriksen, Treves, Moser, & Moser, 2011). Given the fact that theta oscillations are greatly enhanced during working memory tasks, we could imagine that such a mechanism could be also used in prefrontal cortex to test the various possible "context", "rules" or chunks that can be used to describe the current sequence and predict the next events.

6.3 Conclusion

In this thesis, I explored the properties of conscious and unconscious temporal regularity processing. I identified the mismatch negativity as a response representative of unconscious temporal regularity processing. I proposed a neuronal model of the mismatch response based on predictive coding principles that could reproduce the properties of the physiological mismatch response. I showed that the predictive responses where organized in a hierarchical manner in auditory cortex. Finally, I showed that conscious processing using working memory was generating new computational problems and could perform different types of computations that opened new possibilities in terms of temporal regularity learning. Finally, I discussed possible further properties of conscious processing that may arise from the type of encoding that are possible in working memory.

BIBLIOGRAPHY

Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, *14*(8), 1067–1074. doi:10.1038/nn.2869

Ahissar, E., Vaadia, E., Ahissar, M., Bergman, H., Arieli, a, & Abeles, M. (1992). Dependence of cortical plasticity on correlated activity of single neurons and on behavioral context. *Science (New York, N.Y.)*, *257*(5075), 1412–5.

Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Ja, I. P., Lin, F., … Belliveau, J. W. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences*, *101*(17), 6809–6814.

Alain, C., Woods, D. L., & Ogawa, K. H. (1994). Brain indices of automatic pattern processing. *Neuroreport*, *6*(1), 140–4.

Alho, K., Sams, M., Paavilainen, P., Reinikainen, K., & Näätänen, R. (1989). Event-Related Brain Potentials Reflecting Processing of Relevant and Irrelevant Stimuli During Selective Listening. *Psychophysiology*, *26*(5), 514–528. doi:10.1111/j.1469-8986.1989.tb00704.x

Alho, K., Winkler, I., Escera, C., Huotilainen, M., Virtanen, J., Jääskeläinen, I. P., … Ilmoniemi, R. J. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. *Psychophysiology*, *35*(2), 211–24.

Alho, K., Woods, D., & Algazi, A. (1994). Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. *Psychophysiology*, *31*, 469–479.

Allen, J., Kraus, N., & Bradlow, A. (2000). Neural representation of consciously imperceptible speech sound differences. *Attention, Perception, & Psychophysics*, *62*(7), 1383–1393.

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *J Exp Psychol Gen*, *134*(2), 258–81.

Applegate, R. (2007). *Samala-English dictionary : a guide to the Samala language of the Ineseño Chumash People*. Santa Ynez Calif.: Santa Ynez Band of Chumash Indians.

Arcediano, F., Escobar, M., & Miller, R. R. (2005). Bidirectional associations in humans and rats. *Journal of Experimental Psychology. Animal Behavior Processes*, *31*(3), 301–18. doi:10.1037/0097-7403.31.3.301

Asaad, W. F., Rainer, G., & Miller, E. K. (2000). Task-specific neural activity in the primate prefrontal cortex. *Journal of Neurophysiology*, *84*(1), 451–9.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of Conditional Probability Statistics by 8-Month-Old Infants. *Psychological Science*, *9*(4), 321–324. doi:10.1111/1467-9280.00063

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*(3), 183–93.

Baars, B. (1989). *A cognitive theory of consciousness* (Cambridge,.).

Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, *7*(4), 166–172. doi:10.1016/S1364-6613(03)00056-1

Baddeley, A. D., & Hitch, G. (1974). *Working Memory* (pp. 47–89).

Bahlmann, J., Schubotz, R. I., & Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *NeuroImage*, *42*(2), 525–534.

Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D., & Friederici, A. D. (2009). Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Research*, *1298*, 161–170.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron*, *76*(4), 695–711. doi:10.1016/j.neuron.2012.10.038

Bauer, G., Gerstenbrand, F., & Rumpl, E. (1979). Varieties of the locked-in syndrome. *Journal of Neurology*, *221*(2), 77–91.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–21. doi:10.1038/nn1954

Bejjanki, V. R., Beck, J. M., Lu, Z.-L., & Pouget, A. (2011). Perceptual learning as improved probabilistic inference in early sensory areas. *Nature Neuroscience*, *14*(5), 642–8. doi:10.1038/nn.2796

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672–1677.

Bendixen, A., Schröger, E., & Winkler, I. (2009a). I Heard That Coming : Event-Related Potential Evidence for Stimulus-Driven Prediction in the Auditory System. *The Journal of Neuroscience*, *29*(26), 8447– 8451. doi:10.1523/JNEUROSCI.1493-09.2009

Bendixen, A., Schröger, E., & Winkler, I. (2009b). I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(26), 8447–51. doi:10.1523/JNEUROSCI.1493-09.2009

Berger, H. (1929). Uber das Elektrenkephalogramm des Menschen. *Archiv Für Psychiatrie Und Nervenkrankheiten*, *87*(1), 527–570.

Bernat, E., Shevrin, H., & Snodgrass, M. (2001). Subliminal visual oddball stimuli evoke a P300 component. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *112*(1), 159–71.

Bevins, R. (2001). Novelty seeking and reward: Implications for the study of high-risk behaviors. *Current Directions in Psychological Science*, 189–194.

Bevins, R., & Bardo, M. T. (1999). Conditioned increase in place preference by access to novel objects: antagonism by MK-801. *Behavioural Brain Research*, *99*(1), 53–60.

Bi, G., & Poo, M. (1999). Distributed synaptic modification in neural networks induced by patterned stimulation. *Nature*, *401*(6755), 792–6. doi:10.1038/44573

Bi, G., & Poo, M. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Annual Review of Neuroscience*, *24*, 139–166.

Bi, G.-Q., & Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *18*(24), 10464–72.

Blodgett, H. C. (1929). The effect of the introduction of reward upon the maze performance of rats. *University of California Publications in Psychology*, *4*, 113–134.

Bloomfield, T. C., Gentner, T. Q., & Margoliash, D. (2011). What birds have to say about language. *Nature Neuroscience*, *14*(8), 947–948. doi:10.1038/nn.2884

Bor, D., Duncan, J. S., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, *37*(2), 361–7.

Botvinick, M. M., Niv, Y., & Barto, A. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, *113*(3), 262–280. doi:10.1016/j.cognition.2008.08.011

Broadbent, D. (1957). A mechanical model for human attention and immediate memory. *Psychological Review*, *64*(3), 205–215.

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, *63*(1), 119–26. doi:10.1016/j.neuron.2009.06.009

Brosch, M., & Schreiner, C. E. (2000). Sequence sensitivity of neurons in cat primary auditory cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, *10*(12), 1155–67.

Brunel, N., & Wang, X.-J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience*, *11*, 63–85.

Buchler, J. (1955). *Philosophical writings of Peirce.* Oxford, England: Dover.

Buonomano, D. V. (2003). Timing of neural responses in cortical organotypic slices. *Proceedings of the National Academy of Sciences*, *100*(8), 4897–4902.

Buonomano, D. V. (2005). A learning rule for the emergence of stable dynamics and timing in recurrent. *Journal of Neurophysiology*, *94*(4), 2275–2283.

Buonomano, D. V, & Laje, R. (2010). Population clocks: motor timing with neural dynamics. *Trends in Cognitive Sciences*, *14*(12), 520–7. doi:10.1016/j.tics.2010.09.002

Buonomano, D. V, & Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annual Review of Neuroscience*, *21*, 149–86. doi:10.1146/annurev.neuro.21.1.149

Buschman, T. J. T., Siegel, M., Roy, J. E. J. E. J. E., & Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Sciences*, *108*(27), 11252–11255. doi:10.1073/pnas.1104666108/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1104666108

Calford, M. B. (2002). Dynamic representational plasticity in sensory cortex. *Neuroscience*, *111*(4), 709–738.

Camperi, M., & Wang, X. J. (1998). A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. *Journal of Computational Neuroscience*, *5*(4), 383–405.

Carrillo, M. C., Gabrieli, J. D. E., & Schaaf, V. (2000). Selective effects of division of, *28*(3), 293–302.

Carter, R. M., Hofstotter, C., Tsuchiya, N., & Koch, C. (2003). Working memory and fear conditioning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(3), 1399–404. doi:10.1073/pnas.0334049100

Chapman, R. M., & Bragdon, H. R. (1964). Evoked Responses to Numerical and Non-Numerical Visual Stimuli while Problem Solving. *Nature*, *203*(4950), 1155–1157. doi:10.1038/2031155a0

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, *10*(7), 335–44. doi:10.1016/j.tics.2006.05.006

Chimoto, S., Kitama, T., Qin, L., Sakayori, S., & Sato, Y. (2002). Tonal response patterns of primary auditory cortex neurons in alert cats. *Brain Research*, *934*, 34–42.

Chomsky, N. (1956). Three models for the description of language. *Information Theory, IRE Transactions on*.

Chomsky, N. (1957). *Synatctic Structures*.

Clark, R. E. (1998). Classical Conditioning and Brain Systems: The Role of Awareness. *Science*, *280*(5360), 77–81. doi:10.1126/science.280.5360.77

Clark, R. E., Manns, J. R., & Squire, L. R. (2002). Classical conditioning, awareness, and brain systems. *Trends in Cognitive Sciences*, *6*(12), 524–531.

Cohen, D., & Cuffin, B. N. (1983). Demonstration of useful differences between magnetoencephalogram and electroencephalogram. *Electroencephalography and Clinical Neurophysiology*, *56*(1), 38–51. doi:10.1016/0013-4694(83)90005-6

Condon, C. D., & Weinberger, N. M. (1991). Habituation produces frequency-specific plasticity of receptive fields in the auditory cortex. *Behavioral Neuroscience*, *105*(3), 416–30.

Conway, A. R. a., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552. doi:10.1016/j.tics.2003.10.005

Costa-Faidella, J., Grimm, S., Slabu, L., Díaz-Santaella, F., & Escera, C. (2011). Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. *Psychophysiology*, *48*(6), 774–83. doi:10.1111/j.1469-8986.2010.01144.x

Cote, K. a, & Campbell, K. B. (1999). P300 to high intensity stimuli during REM sleep. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *110*(8), 1345–50.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. doi:10.1017/S0140525X01003922

Cowan, N. (2010). The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, *19*(1), 51–57. doi:10.1177/0963721409359277

Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 909.

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. a, Fernández-Espejo, D., Pickard, J. D., … Owen, A. M. (2011). Bedside detection of awareness in the vegetative state: a cohort study. *Lancet*, *378*(9809), 2088–94. doi:10.1016/S0140-6736(11)61224-5

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. a, Fernández-Espejo, D., Pickard, J. D., … Owen, A. M. (2013). Reanalysis of "Bedside detection of awareness in the vegetative state: a cohort study" - Authors' reply. *Lancet*, *381*(9863), 291–2. doi:10.1016/S0140-6736(13)60126-9

Culy, C. (1985). The complexity of the vocabulary of Bambara. *Linguistics and Philosophy*, *8*(3), 345–351.

Curran, T., & Keele, S. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*(1), 189–202.

D'Esposito, M., Postle, B., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain and Cognition*, *41*(1), 66–86. doi:10.1006/brcg.1999.1096

Daum, I., Schugens, M. M., Ackermann, H., Lutzenberger, W., Dichgans, J., & Birbaumer, N. (1993). Classical conditioning after cerebellar lesions in humans. *Behavioral Neuroscience*, *107*(5), 748–56.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–11. doi:10.1038/nn1560

Dayan, P. (2007). Bilinearity, rules, and prefrontal cortex. *Frontiers in Computational Neuroscience*, *1*(November), 1. doi:10.3389/neuro.10.001.2007

Dayan, P. (2012). How to set the switches on this thing. *Current Opinion in Neurobiology*, *22*(6), 1068–1074. doi:10.1016/j.conb.2012.05.011

Dayan, P., & Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Current Opinion in Neurobiology*, *18*(2), 185–96. doi:10.1016/j.conb.2008.08.003

Deacon, T. (1997). *The symbolic species. The co-evolution of language and the human brain.*

Debanne, D., Shulz, D. E., & Frégnac, Y. (1998). Activity-dependent regulation of "on" and "off" responses in cat visual cortical receptive fields. *The Journal of Physiology*, *508 (Pt 2*, 523–48.

Dehaene, S. (2008). Conscious and nonconscious processes: distinct forms of evidence accumulation? In C. Engel & W. Singer (Eds.), *Better Than Conscious? Decision Making, the Human Mind, and Implications For Institutions. Strüngmann Forum Report.* Cambridge: MIT Press.

Dehaene, S., & Changeux, J.-P. (1991). The Wisconsin Card Sorting Test: Theoretical analysis and modeling in a neuronal network. *Cerebral Cortex.*

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*(2), 200–27. doi:10.1016/j.neuron.2011.03.018

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211.

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998a). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, *95*(24), 14529–14534.

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998b). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(24), 14529–34.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*, 1–37.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, *8*(4), 919.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology*, *5*(10), e260. doi:10.1371/journal.pbio.0050260

Den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2009). A dual role for prediction error in associative learning. *Cerebral Cortex*, *19*(5), 1175–85. doi:10.1093/cercor/bhn161

Dennett, D. (1992). *Consciousness Explained*. Back Bay Books.

Deouell, L. Y., Parnes, A., Pickard, N., & Knight, R. T. (2006). Spatial location is accurately tracked by human auditory sensory memory: evidence from the mismatch negativity. *The European Journal of Neuroscience*, *24*(5), 1488–94. doi:10.1111/j.1460-9568.2006.05025.x

Desmedt, J., Debrecker, J., & Manil, J. (1965). Mise en évidence d'un signe électrique cérébral associé à la détection par le sujet d'un stimulus sensoriel tactile. *Bull Acad R Med Belg*, *5*, 887–936.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*(2), 343–350.

Dezfouli, A., & Balleine, B. W. (2013). Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Computational Biology*, *9*(12), e1003364. doi:10.1371/journal.pcbi.1003364

DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, *11*(8), 333–41. doi:10.1016/j.tics.2007.06.010

Diekhof, E. K., Biedermann, F., Ruebsamen, R., & Gruber, O. (2009). Top-down and bottom-up modulation of brain structures involved in auditory discrimination. *Brain Research*, *1297*, 118–23. doi:10.1016/j.brainres.2009.08.040

Dienes, Z., & Altmann, G. (1997). Transfer of implicit knowledge across domains: How implicit and how abstract. *How Implicit Is Implicit Learning*, *5*, 107–123.

Dominey, P., Arbib, M., & Joseph, J. P. (1995). A model of corticostriatal plasticity for learning oculomotor associations and sequences. *Journal of Cognitive Neuroscience*, *7*(3), 311–36. doi:10.1162/jocn.1995.7.3.311

Dominey, P. F., Lelekov, T., Ventre-Dominey, J., & Jeannerod, M. (1998). Dissociable processes for learning the surface structure and abstract structure of sensorimotor sequences. *Journal of Cognitive Neuroscience*, *10*(6), 734–51.

Dong, D., & Atick, J. (1995). Statistics of natural time-varying images. *Network: Computation in Neural Systems*, *6*(3), 345–358. doi:10.1088/0954-898X/6/3/003

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks : The Official Journal of the International Neural Network Society*, *12*(7-8), 961–974.

Durstewitz, D., Kelc, M., & Güntürkün, O. (1999). A neurocomputational theory of the dopaminergic modulation of working memory functions. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(7), 2807–22.

Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000). Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *Journal of Neurophysiology*, *83*(3), 1733–50.

Edelman, G. M. (1989). *The remembered present: a biological theory of consciousness* (p. 346).

Ehrlichman, R. S., Maxwell, C. R., Majumdar, S., & Siegel, S. J. (2008). Deviance-elicited changes in event-related potentials are attenuated by ketamine in mice. *Journal of Cognitive Neuroscience*, *20*(8), 1403–14. doi:10.1162/jocn.2008.20097

Elman, J. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 71–99.

Endress, A. D., Dehaene-Lambertz, G., & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition*, *105*(3), 577–614. doi:10.1016/j.cognition.2006.12.014

Endress, A. D., & Wood, J. (2011). From movements to actions: Two mechanisms for learning action sequences. *Cognitive Psychology*, *63*(3), 141–171. doi:10.1016/j.cogpsych.2011.07.001

Engel, T. a., & Wang, X.-J. (2011). Same or Different? A Neural Circuit Mechanism of Similarity-Based Pattern Match Decision Making. *Journal of Neuroscience*, *31*(19), 6982–6996. doi:10.1523/JNEUROSCI.6150-10.2011

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

Faugeras, F., Rohaut, B., Weiss, N., Bekinschtein, T. A., Galanaud, D., Puybasset, L., … Naccache, L. (2011). Probing consciousness with event-related potentials in patients who meet clinical criteria for vegetative state. *Neurology*, *(in press)*.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, *4*(December), 215. doi:10.3389/fnhum.2010.00215

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, *1*(1), 1–47.

Fiorillo, C. D. (2008). Towards a general theory of neural computation based on prediction by single neurons. *PloS One*, *3*(10), e3298. doi:10.1371/journal.pone.0003298

Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: an overview. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1598), 1933–55. doi:10.1098/rstb.2012.0103

Ford, J. M., & Hillyard, S. A. (1981). Event-Related Potentials (ERP s) to Interruptions of a Steady Rhythm. *Psychophysiology*, *18*(3), 322–330. doi:10.1111/j.1469-8986.1981.tb03043.x

Frank, M. C., & Tenenbaum, J. B. (2011). Three ideal observer models for rule learning in simple languages. *Cognition*, *120*(3), 360–71. doi:10.1016/j.cognition.2010.10.005

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective & Behavioral Neuroscience*, *1*(2), 137–60.

Freedman, D. J., Riesenhuber, M., Poggio, T. A., & Miller, E. K. (2002). Visual categorization and the primate prefrontal cortex: neurophysiology and behavior. *Journal of Neurophysiology*, *88*(2), 929–41.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science (New York, N.Y.)*, *291*(5502), 312–6. doi:10.1126/science.291.5502.312

Frégnac, Y., Shulz, D. E., Thorpe, S. J., & Bienenstock, E. (1992). Cellular Analogs of Visual Cortical Epigenesis . Orientation Selectivity I . Plasticity of orientation selection. *Journal of Neuroscience*, *12*(4), 1280–1300.

Friston, K. J. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B.*, *360*, 815–836. doi:10.1098/rstb.2005.1622

Friston, K. J., & Kiebel, S. (2009a). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1211–21. doi:10.1098/rstb.2008.0300

Friston, K. J., & Kiebel, S. (2009b). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1211–21. doi:10.1098/rstb.2008.0300

Friston, K. J., Kilner, J. M., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology*, *100*, 70–87. doi:10.1016/j.jphysparis.2006.10.001

Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, *159*(3), 417–458. doi:10.1007/s11229-007-9237-y

Fuentemilla, L., Penny, W. D., Cashdollar, N., Bunzeck, N., & Düzel, E. (2010). Theta-coupled periodic replay in working memory. *Current Biology : CB*, *20*(7), 606–12. doi:10.1016/j.cub.2010.01.057

Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, *20*(2), 177–82. doi:10.1016/j.conb.2010.03.005

Fuster, J. M. (1971). Neuron activity related to short-term memory. *Science*, *173*(3997), 652–654.

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences*, *101*(36), 13124–31. doi:10.1073/pnas.0404965101

Gao, W. J., Krimer, L. S., & Goldman-Rakic, P. S. (2001). Presynaptic regulation of recurrent excitation by D1 receptors in prefrontal circuits. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(1), 295–300. doi:10.1073/pnas.011524298

Garrido, M. I., Friston, K. J., Kiebel, S. J., Stephan, K. E., Baldeweg, T., & Kilner, J. M. (2008). The functional anatomy of the MMN: a DCM study of the roving paradigm. *NeuroImage*, *42*(2), 936–44. doi:10.1016/j.neuroimage.2008.05.018

Garrido, M. I., Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2007). Evoked brain responses are generated by feedback loops. *Proceedings of the National Academy of Sciences*, *104*(52), 20961–20966.

Garrido, M. I., Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2009). Dynamic causal modeling of the response to frequency deviants. *Journal of Neurophysiology*, *101*(5), 2620–2631. doi:10.1152/jn.90291.2008

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). Clinical Neurophysiology The mismatch negativity : A review of underlying mechanisms. *Clinical Neurophysiology*, *120*(3), 453–463. doi:10.1016/j.clinph.2008.11.029

Gauthier, B., Eger, E., Hesselmann, G., Giraud, A.-L., & Kleinschmidt, A. (2012). Temporal tuning properties along the human ventral visual stream. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(41), 14433–41. doi:10.1523/JNEUROSCI.2467-12.2012

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

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, *440*(April), 1204–1207. doi:10.1038/nature04675

Gerken, L. (2006). Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition*, *98*(3), B67–74. doi:10.1016/j.cognition.2005.03.003

Giard, M. H., Lavikahen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., & Näätänen, R. (1995). Separate Representation of Stimulus Frequency, Intensity, and Duration in Auditory Sensory Memory: An Event-Related Potential and Dipole-Model Analysis. *Journal of Cognitive Neuroscience*, *7*(2), 133–143. doi:10.1162/jocn.1995.7.2.133

Gilmartin, M. R., & Helmstetter, F. J. (2010). Trace and contextual fear conditioning require neural activity and NMDA receptor-dependent transmission in the medial prefrontal cortex. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *17*(6), 289–96. doi:10.1101/lm.1597410

Gilmartin, M. R., Miyawaki, H., Helmstetter, F. J., & Diba, K. (2013). Prefrontal activity links nonoverlapping events in memory. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(26), 10910–4. doi:10.1523/JNEUROSCI.0144-13.2013

Gläscher, J., Daw, N. D., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–95. doi:10.1016/j.neuron.2010.04.016

Goldfine, A., Bardin, J., Noirhomme, Q., Fins, J. J., Schiff, N. D., & Victor, J. D. (2013). Reanalysis of "Bedside detection of awareness in the vegetative state: a cohort study." *Lancet*, *381*(9863), 289–291. doi:10.1016/S0140-6736(13)60125-7.Reanalysis

Goldfine, A. M., Victor, J. D., Conte, M. M., Bardin, J. C., & Schiff, N. D. (2011). Determination of awareness in patients with severe brain injury using EEG power spectral analysis. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *122*(11), 2157–68. doi:10.1016/j.clinph.2011.03.022

Goldman, M. S. (2009). Memory without feedback in a neural network. *Neuron*, *61*(4), 621–34. doi:10.1016/j.neuron.2008.12.012

Goldman-rakic, P. S. (1995). Cellular Basis of Working Memory. *Neuron*, *14*, 477–485.

Gomes, H., Bernstein, R., Ritter, W., Vaughan, H. G., & Miller, J. (1997). Storage of feature conjunctions in transient auditory memory. *Psychophysiology*, *34*(6), 712–6.

Gómez, R. L. (2002). Variability and detection of invariant strucutre. *Psychological Science*, *13*(5), 431–436.

Graybiel, A. (2008). Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* doi:10.1146/annurev.neuro.29.051605.112851

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory*, *70*(1-2), 119–36. doi:10.1006/nlme.1998.3843

Grossberg, S. (1970). Some networks that can learn, remember, and reproduce any number of complicated space-time patterns. I. *Studies in Applied Mathematics*, *49*(2), 135–166.

Gruber, A. J., Dayan, P., Gutkin, B. S., & Solla, S. a. (2006a). Dopamine modulation in the basal ganglia locks the gate to working memory. *Journal of Computational Neuroscience*, *20*(2), 153–66. doi:10.1007/s10827-005-5705-x

Gruber, A. J., Dayan, P., Gutkin, B. S., & Solla, S. a. (2006b). Dopamine modulation in the basal ganglia locks the gate to working memory. *Journal of Computational Neuroscience*, *20*(2), 153–66. doi:10.1007/s10827-005-5705-x

Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biology*, *6*(6), e138. doi:10.1371/journal.pbio.0060138

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *25*(45), 10494–501. doi:10.1523/JNEUROSCI.1227-05.2005

Haider, H., Eichler, A., & Lange, T. (2011). An old problem: How can we distinguish between conscious and unconscious knowledge acquired in an implicit learning task? *Consciousness and Cognition*, *20*(3), 658–72. doi:10.1016/j.concog.2010.10.021

Haider, H., & Frensch, P. a. (2005). The generation of conscious awareness in an incidental learning situation. *Psychological Research*, *69*(5-6), 399–411. doi:10.1007/s00426-004-0209-2

Haider, H., & Frensch, P. a. (2009). Conflicts between expected and actually performed behavior lead to verbal report of incidentally acquired sequential knowledge. *Psychological Research*, *73*(6), 817–34. doi:10.1007/s00426-008-0199-6

Hakes, D. T., Evans, J. S., & Brannon, L. L. (1976). Understanding sentences with relative clauses. *Memory & Cognition*, *4*(3), 283–90. doi:10.3758/BF03213177

Harrison, S. S. a, & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632–635. doi:10.1038/nature07832.Decoding

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, *298*(5598), 1569–79. doi:10.1126/science.298.5598.1569

Hayden, B. Y. (2010). Ambiguity aversion in rhesus macaques. *Frontiers in Neuroscience*, *4*(September), 1–7. doi:10.3389/fnins.2010.00166

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory*. (Wiley & so.). New York.

Heinke, W., Kenntner, R., Gunter, T. C., Sammler, D., Olthoff, D., & Koelsch, S. (2004). Sequential effects of increasing propofol sedation on frontal and temporal cortices as indexed by auditory event-related potentials. *Anesthesiology*, *100*(3), 617–25.

Heinz, J., & Idsardi, W. (2011). Psychology. Sentence and word complexity. *Science (New York, N.Y.)*, *333*(6040), 295–7. doi:10.1126/science.1210358

Helmholtz, H. (1860). *Handbuch der physiologischen Optik* (Dover.). New york: (English trans., Southall JPC, Ed.).

Herd, S., Mingus, B., & O'Reilly, R. C. (2010). Dopamine and self-directed learning. *Biologically Inspired Cognitive Architectures 2010: Proceedings of the First Annual Meeting of the BICA Society*, 58–63.

Hillyard, S. A., Squires, K. C., Bauer, J. W., Lindsay, P. H., Carroll, B. J., & Sharp, P. T. (1971). Evoked Potential Correlates of Auditory Signal Detection. *Science*, *172*(3990), 1357–1360. doi:10.1126/science.172.3990.1357

Hinton, G. E. (2007). Learning multiple layers of representation. *Trends in Cognitive Sciences*, *11*(10), 428–34. doi:10.1016/j.tics.2007.09.004

Holland, P. C., & Rescorla, R. A. (1975). Second-order conditioning with food unconditioned stimulus. *Journal of Comparative and Physiological Psychology*, *88*(1), 459–67.

Horváth, J., Czigler, I., Sussman, E., & Winkler, I. (2001). Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. *Cognitive Brain Research*, *12*, 131–144.

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

Horváth, J., Winkler, I., & Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biological Psychology*, *79*(2), 139–47. doi:10.1016/j.biopsycho.2008.04.001

Hoshi, E., Shima, K., & Tanji, J. (1998). Task-dependent selectivity of movement-related neuronal activity in the primate prefrontal cortex. *Journal of Neurophysiology*, 3392–3397.

Hosoya, T., Baccus, S. a, & Meister, M. (2005). Dynamic predictive coding by the retina. *Nature*, *436*(7047), 71–7. doi:10.1038/nature03689

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 580–593. doi:10.1002/wcs.142

Hubel, D., & Wiesel, T. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*, 215–243.

Huettel, S., Mack, P. B., & McCarthy, G. (2002). Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. *Nature Neuroscience*, *5*(5), 485–90. doi:10.1038/nn841

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

Huybregts, R. (1984). The weak inadequacy of context-free phrase structure grammars. In *Van periferie naar kern* (pp. 81–99).

Ivry, R., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, *1*(2), 136–52. doi:10.1162/jocn.1989.1.2.136

Izhikevich, E. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, *14*(6), 1569–1572.

Jacobsen, T., & Schröger, E. (2003). Measuring duration mismatch negativity. *Clinical Neurophysiology*, *114*(6), 1133–1143. doi:10.1016/S1388-2457(03)00043-9

Jaeger, H. (2001). The "'echo state"' approach to analysing and training recurrent neural networks. *GMD Report 148. GMD — German National Research Institute for Computer Science.*

Jäger, G., & Rogers, J. (2012). Formal language theory: refining the Chomsky hierarchy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1598), 1956–70. doi:10.1098/rstb.2012.0077

Javitt, D. C., Steinschneider, M., Schroeder, C. E., & Arezzo, J. (1996). Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. *Proceedings of the National Academy of Sciences*, *93*, 11962–11967.

Jehee, J. F. M., Rothkopf, C., Beck, J. M., & Ballard, D. H. (2006). Learning receptive fields using predictive feedback. *Journal of Physiology, Paris*, *100*(1-3), 125–32. doi:10.1016/j.jphysparis.2006.09.011

Jennett, B., & Plum, F. (1972). Persistent vegetative state after brain damage. *The Lancet*, 734–737.

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *Neuroscience*, *15*, 1395–1399.

Jezek, K., Henriksen, E. J., Treves, A., Moser, E. I., & Moser, M.-B. (2011). Theta-paced flickering between place-cell maps in the hippocampus. *Nature*, *478*(7368), 246–9. doi:10.1038/nature10439

Joutsiniemi, S.-L., & Hari, R. (1989). Omissions of Auditory Stimuli May Activate Frontal Cortex. *The European Journal of Neuroscience*, *1*(5), 524–528.

Kalmbach, B. E., Ohyama, T., & Mauk, M. D. (2010). Temporal patterns of inputs to cerebellum necessary and sufficient for trace eyelid conditioning. *Journal of Neurophysiology*, *104*(2), 627–40. doi:10.1152/jn.00169.2010

Karlsson, F. (2007). Constraints on multiple center-embedding of clauses. *JOURNAL OF LINGUISTICS-CAMBRIDGE-*, *43*(December 2006), 365–392.

Kehoe, E. J., Graham-Clarke, P., & Schreurs, B. G. (1989). Temporal patterns of the rabbit's nictitating membrane response to compound and component stimuli under mixed CS–US intervals. *Behavioral Neuroscience*, *103*(2), 283–295.

Kekoni, J., Hamalainen, H., Saarinen, M., Grohn, J., Reinikainen, K., Lehtokoski, A., & Näätänen, R. (1997). Rate effect and mismatch responses in the somatosensory system: ERP-recordings in humans. *Biological Psychology*, *46*(2), 125–142.

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(31), 10687–92. doi:10.1073/pnas.0802631105

Kiebel, S. J., Daunizeau, J., & Friston, K. J. (2008). A Hierarchy of Time-Scales and the Brain. *PLoS Computational Biology*, *4*(11). doi:10.1371/journal.pcbi.1000209

Kiebel, S. J., Daunizeau, J., & Friston, K. J. (2009). Perception and hierarchical dynamics. *Frontiers in Neuroinformatics*, *3*(July), 20.

Kiebel, S. J., Kriegstein, K. Von, Daunizeau, J., & Friston, K. J. (2009). Recognizing Sequences of Sequences. *PLoS Computational Biology*, *5*(8), 1–13. doi:10.1371/journal.pcbi.1000464

Kim, J. J., Clark, R. E., & Thompson, R. F. (1995). Hippocampectomy impairs the memory of recently, but not remotely, acquired trace eyeblink conditioned responses. *Behavioral Neuroscience*, *109*(2), 195–203.

Kinder, a, & Assmann, a. (2000). Learning artificial grammars: no evidence for the acquisition of rules. *Memory & Cognition*, *28*(8), 1321–32.

Knight, R. (1996). Contribution of human hippocampal region to novelty detection. *Nature*, *383*(6597), 256–9. doi:10.1038/383256a0

Knight, R. T. (1984). Decreased response to novel stimuli after prefrontal lesions in man. *Electroencephalography and Clinical Neurophysiology*, *59*(1), 9–20.

Knight, R. T. (1997). Distributed cortical network for visual attention. *Journal of Cognitive Neuroscience*, *9*(1), 75–91.

Kobayashi, S., & Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*, 7837–7846. doi:10.1523/JNEUROSCI.1600-08.2008

Koelsch, S., Heinke, W., Sammler, D., & Olthoff, D. (2006). Auditory processing during deep propofol sedation and recovery from unconsciousness. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *117*(8), 1746–59. doi:10.1016/j.clinph.2006.05.009

Korzyukov, O. A., Winkler, I., Gumenyuk, V., & Alho, K. (2003). Processing abstract auditory features in the human auditory cortex. *Neuroimage*, *20*, 2245–2258. doi:10.1016/j.neuroimage.2003.08.014

Krauel, K., Schott, P., Sojka, B., Pause, B. M., & Ferstl, R. (1999). Is There a Mismatch Negativity Analogue in the Olfactory Event-Related Potential? *Journal of Psychophysiology*, *13*(1), 49–55. doi:10.1027//0269-8803.13.1.49

Kronforst-Collins, M. a, & Disterhoft, J. F. (1998). Lesions of the caudal area of rabbit medial prefrontal cortex impair trace eyeblink conditioning. *Neurobiology of Learning and Memory*, *69*(2), 147–62. doi:10.1006/nlme.1997.3818

Laje, R., & Buonomano, D. V. (2013). Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nature Neuroscience*, *16*(7), 925–933. doi:10.1038/nn.3405

Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, *1*(3), 204–20. doi:10.1080/17588921003731586

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–9.

Lamy, D., Salti, M., & Bar-Haim, Y. (2009). Neural correlates of subjective awareness and unconscious processing: an ERP study. *Journal of Cognitive Neuroscience*, *21*(7), 1435–46. doi:10.1162/jocn.2009.21064

Lange, K. W., Robbins, T. W., Marsden, C. D., James, M., Owen, A. M., & Paup, G. M. (1992). L-Dopa withdrawal in Parkinson ' s disease selectively impairs cognitive performance in tests sensitive to frontal lobe dysfunction. *Psychopharmacology*, *107*, 394–404.

Lau, H. C. (2008). A higher order Bayesian decision theory of consciousness. *Progress in Brain Research*, *168*(07), 35–48. doi:10.1016/S0079-6123(07)68004-2

Lee, H., Simpson, G., & Logothetis, N. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron*, *45*, 147–156. doi:10.1016/j.neuron.2004.12.025

Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nature Neuroscience*, *16*(8), 997–999. doi:10.1038/nn.3452

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *20*(7), 1434–48.

Levy, W., & Steward, O. (1983). Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience*, *8*(4), 791–797.

Lewicki, M. S. (2002). Efficient coding of natural sounds. *Nature Neuroscience*, *5*(4), 356–63. doi:10.1038/nn831

Lewis, S. J. G., Dove, A., Robbins, T. W., Barker, R. A., & Owen, A. M. (2004). Striatal contributions to working memory: a functional magnetic resonance imaging study in humans. *European Journal of Neuroscience*, *19*(3), 755–760. doi:10.1111/j.1460-9568.2004.03108.x

Linden, D. E. J., Oosterhof, N. N., Klein, C., & Downing, P. E. (2012). Mapping brain activation and information during category-specific visual working memory. *Journal of Neurophysiology*, *107*(2), 628–39. doi:10.1152/jn.00105.2011

Lisman, J. E., & Idiart, M. A. P. (1995). Storage of 7±2 short-term memories in oscillatory subcycles. *Science*, *267*(5203), 1512–1515.

Locke, J. (1690). *An essay concerning human understanding.*

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *193*(1996), 1996–1998.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, 1–10. doi:10.1016/j.tics.2013.06.006

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. doi:10.1038/nn.3655

Maass, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Computation*, *14*(11), 2531–60. doi:10.1162/089976602760407955

Maass, W., Natschlager, T., Markram, H., & Natschläger, T. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Computation*, *14*(11), 2531–60. doi:10.1162/089976602760407955

Macar, F., & Vidal, F. (2003). The CNV peak: an index of decision making and temporal memory. *Psychophysiology*, *40*(6), 950–4.

Macar, F., & Vidal, F. (2004). Event-Related Potentials as Indices of Time Processing : A Review. *Journal of Psychophysiology*, *18*, 89–104. doi:10.1027/0269-8803.18.2

Macar, F., Vidal, F., & Casini, L. (1999). The supplementary motor area in motor and sensory timing: evidence from slow brain potential changes. *Experimental Brain Research*, *125*, 271–280.

Machens, C. K., Romo, R., & Brody, C. D. (2010). Functional , But Not Anatomical , Separation of " What " and " When " in Prefrontal Cortex. *Cortex*, *30*(1), 350 –360. doi:10.1523/JNEUROSCI.3276-09.2010

Manns, J. R., Clark, R. E., & Squire, L. R. (2002). Standard delay eyeblink classical conditioning is independent of awareness. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(1), 32–37. doi:10.1037//0097-7403.28.1.32

Mäntysalo, S., & Näätänen, R. (1987). The duration of a neuronal trace of an auditory stimulus as indicated by event-related potentials. *Biological Psychology*, *24*(3), 183–195. doi:10.1016/0301-0511(87)90001-9

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science (New York, N.Y.)*, *283*(5398), 77–80.

Markram, H., Lübke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, *275*(5297), 213–5.

Markram, H., Wang, Y., & Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(9), 5323–8.

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. *Inc., New York, NY*.

Marreiros, A. C., Kiebel, S. J., Daunizeau, J., Harrison, L. M., & Friston, K. J. (2009). Population dynamics under the Laplace assumption. *Neuroimage*, *44*(3), 701–714. doi:10.1016/j.neuroimage.2008.10.008

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(47), 12539–45. doi:10.1523/JNEUROSCI.2925-08.2008

Marti, S., Sigman, M., & Dehaene, S. (2012). A shared cortical bottleneck underlying Attentional Blink and Psychological Refractory Period. *NeuroImage*, *59*(3), 2883–98. doi:10.1016/j.neuroimage.2011.09.063

Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annual Review of Neuroscience*, *27*, 307–40. doi:10.1146/annurev.neuro.27.070203.144247

Mauk, M. D., & Thompson, R. F. (1987). Retention of classically conditioned eyelid responses following acute decerebration. *Brain Research*, *403*(1), 89–95.

May, P., & Tiitinen, H. (2009). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, *47*(1), 66–122. doi:10.1111/j.1469-8986.2009.00856.x

McGlinchey-Berroth, R., Carrillo, M. C., Gabrieli, J. D., Brawn, C. M., & Disterhoft, J. F. (1997). Impaired trace eyeblink conditioning in bilateral, medial-temporal lobe amnesia. *Behavioral Neuroscience*, *111*(5), 873–82.

McLaughlin, J., Skaggs, H., Churchwell, J., & Powell, D. a. (2002). Medial prefrontal cortex and Pavlovian conditioning: Trace versus delay conditioning. *Behavioral Neuroscience*, *116*(1), 37–47. doi:10.1037//0735-7044.116.1.37

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, *11*(1), 103–7. doi:10.1038/nn2024

Menon, V., Anagnoson, R. T., Glover, G. H., & Pfefferbaum, a. (2000). Basal ganglia involvement in memory-guided movement sequencing. *Neuroreport*, *11*(16), 3641–5.

Meyer, T., & Olson, C. R. (2011). Statistical learning of visual transitions in monkey inferotemporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(48). doi:10.1073/pnas.1112895108

Meyer, T., Qi, X.-L., & Constantinidis, C. (2007). Persistent discharges in the prefrontal cortex of monkeys naive to working memory tasks. *Cerebral Cortex (New York, N.Y. : 1991)*, *17 Suppl 1*, i70–6. doi:10.1093/cercor/bhm063

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. doi:10.1146/annurev.neuro.24.1.167

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *The Journal of Neuroscience*, *16*(16), 5154.

Miller, E. K., Freedman, D. J., Wallis, J. D., Trans, P., & Lond, R. S. (2002). The prefrontal cortex: categories, concepts and cognition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *357*(1424), 1123–36. doi:10.1098/rstb.2002.1099

Miller, E. K., Nieder, A., Freedman, D. J., & Wallis, J. D. (2003). Neural correlates of categories and concepts. *Current Opinion in Neurobiology*, *13*(2), 198–203. doi:10.1016/S0959-4388(03)00037-0

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *101*(2), 343–52.

Miyoshi, E., Wietzikoski, S., Camplessei, M., Silveira, R., Takahashi, R. N., & Da Cunha, C. (2002). Impaired learning in a spatial working memory version and in a cued version of the water maze in rats with MPTP-induced mesencephalic dopaminergic lesions. *Brain Research Bulletin*, *58*(1), 41–7.

Moerel, M., De Martino, F., & Formisano, E. (2012). Processing of natural sounds in human auditory cortex: tonotopy, spectral tuning, and relation to voice sensitivity. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(41), 14205–16. doi:10.1523/JNEUROSCI.1388-12.2012

Mongillo, G., Amit, D. J., & Brunel, N. (2003). Retrospective and prospective persistent activity induced by Hebbian learning in a recurrent cortical network. *European Journal of Neuroscience*, *18*(7), 2011–2024. doi:10.1046/j.1460-9568.2003.02908.x

Monti, M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L. J.-F., … Laureys, S. (2010). Willful modulation of brain activity in disorders of consciousness. *New England Journal of Medicine*, *362*(7), 579–589.

Moyer, J. R., Deyo, R. A., & Disterhoft, J. F. (1990). Hippocampectomy disrupts trace eye-blink conditioning in rabbits. *Behavioral Neuroscience*, *104*(2), 243–52.

Müller, U., von Cramon, D. Y., & Pollmann, S. (1998). D1- versus D2-receptor modulation of visuospatial working memory in humans. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *18*(7), 2720–8.

Muller-Gass, A., Stelmack, R. M., & Campbell, K. B. (2005). "...and Were Instructed To Read a Self-Selected Book While Ignoring the Auditory Stimuli": the Effects of Task Demands on the Mismatch Negativity. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *116*(9), 2142–52. doi:10.1016/j.clinph.2005.05.012

Mushiake, H., Saito, N., Sakamoto, K., Itoyama, Y., & Tanji, J. (2006). Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. *Neuron*, *50*(4), 631–641. doi:10.1016/j.neuron.2006.03.045

Näätänen, R. (2003). Mismatch negativity: clinical research and possible applications. *International Journal of Psychophysiology*, *48*(2), 179–188. doi:10.1016/S0167-8760(03)00053-9

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*(4), 313–29.

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

Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1987). The mismatch negativity to intensity changes in an auditory stimulus sequence. *Electroencephalography and Clinical Neurophysiology*, *40*, 125–31.

Näätänen, R., Paavilainen, P., & Reinikainen, K. (1989). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neuroscience Letters*, *107*(1-3), 347–52.

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

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). "Primitive intelligence" in the auditory cortex. *Trends in Neurosciences*, *24*(5), 283–8.

Nieder, A. (2009). Prefrontal cortex and the evolution of symbolic reference. *Current Opinion in Neurobiology*, *19*(1), 99–108. doi:10.1016/j.conb.2009.04.008

Ninokura, Y., Mushiake, H., & Tanji, J. (2003). Representation of the temporal order of visual objects in the primate lateral prefrontal cortex. *Journal of Neurophysiology*, *89*(5), 2868–73. doi:10.1152/jn.00647.2002

Ninokura, Y., Mushiake, H., & Tanji, J. (2004). Integration of temporal order and object information in the monkey lateral prefrontal cortex. *Journal of Neurophysiology*, *91*(1), 555–60. doi:10.1152/jn.00694.2003

Nissen, M., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1–32.

Norman, D., & Shallice, T. (1986). Attention to action: willed and automatic control of behavior.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*(2), 283–328. doi:10.1162/089976606775093909

Offen, S., Schluppeck, D., & Heeger, D. J. (2009). The role of early visual cortex in visual short-term memory and visual attention. *Vision Research*, *49*(10), 1352–62. doi:10.1016/j.visres.2007.12.022

Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(24), 8776–80. doi:10.1073/pnas.0500810102

Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(7), 2745–50. doi:10.1073/pnas.0708424105

Orlov, T., Yakovlev, V., Hochstein, S., & Zohary, E. (2000). Macaque monkeys categorize images by their ordinal number. *Nature*, *404*(6773), 77–80. doi:10.1038/35003571

Otto, A., Gershman, S., Markman, A. B., & Daw, N. D. (2013). The Curse of Planning Dissecting Multiple Reinforcement-Learning Systems by Taxing the Central Executive. *Psychological Science*, *24*(5), 751–761.

Oudeyer, P.-Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *Evolutionary Computation, IEEE Transactions on*, *11*(2), 265–286.

Owen, a. M., James, M., Leigh, P. N., Summers, B. a., Marsden, C. D., Quinn, N. P., … Robbins, T. W. (1992). Fronto-Striatal Cognitive Deficits At Different Stages of Parkinson'S Disease. *Brain*, *115*(6), 1727–1751. doi:10.1093/brain/115.6.1727

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science (New York, N.Y.)*, *313*(5792), 1402. doi:10.1126/science.1130197

Paavilainen, P., Jaramillo, M., Näätänen, R., & Winkler, I. (1999). Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neuroscience Letters*, *265*, 179–182.

Parmentier, F. B. R., Elsley, J. V, Andrés, P., & Barceló, F. (2011). Why are auditory novels distracting? Contrasting the roles of novelty, violation of expectation and stimulus change. *Cognition*, *119*(3), 374–380. doi:10.1016/j.cognition.2011.02.001

Pascanu, R., & Jaeger, H. (2010). A neurodynamical model for working memory. *Neural Networks*, *24*(2), 199–207. doi:10.1016/j.neunet.2010.10.003

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews. Neuroscience*, *6*(2), 97–107. doi:10.1038/nrn1603

Pause, B. M., & Krauel, K. (2000). Chemosensory event-related potentials (CSERP) as a key to the psychology of odors. *International Journal of Psychophysiology*, *36*(2), 105–22.

Pavlov, I. (1927). *Conditioned reflexes. An investigation of the physiological activity of the cerebral coortex.* (H. Milford, Ed.) (Oxford Uni.).

Pazo-alvarez, P., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: a review. *Biological Psychology*, *63*, 199–236.

Pegado, F., Bekinschtein, T. A., Chausson, N., Dehaene, S., Cohen, L., & Naccache, L. (2010). Probing the lifetimes of auditory novelty detection processes. *Neuropsychologia*, *48*(10), 3145–54. doi:10.1016/j.neuropsychologia.2010.06.030

Perfors, A., Tenenbaum, J., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, *118*, 306–338.

Perrin, F., Garcia-Larrea, L., Mauguière, F., & Bastuji, H. (1999). A differential brain response to the subject ' s own name persists during sleep. *Clinical Neurophysiology*, *110*, 2153–2164.

Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, *10*(2), 257–61. doi:10.1038/nn1840

Petersson, K. M., & Hagoort, P. (2012). The neurobiology of syntax: beyond string sets. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1598), 1971–83. doi:10.1098/rstb.2012.0101

Plum, F., & Posner, J. B. (1966). *The Diagnosis of Stupor and Coma* (p. 197).

Polich, J. (2007). Updating P300 : An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, 2128–2148. doi:10.1016/j.clinph.2007.04.019

Polich, J., & Criado, J. R. (2006). Neuropsychology and neurophannacology of P3a and P3b. *International Journal of Psychophysiology*, *60*(2), 172–185.

Posner, M. I., & Rothbart, M. K. (1998). Attention, self-regulation and consciousness. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *353*(1377), 1915–27. doi:10.1098/rstb.1998.0344

Postle, B., Druzgal, T., & D'Esposito, M. (2003). Seeking the neural substrates of visual working memory storage. *Cortex*, *39*(4-5), 927–946.

Pouget, A., Deneve, S., & Duhamel, J. (2002). spatial representations, *3*(September), 1–7.

Powell, D. A., Skaggs, H., Churchwell, J., & McLaughlin, J. (2001). Posttraining lesions of the medial prefrontal cortex impair performance of Pavlovian eyeblink conditioning but have no effect on concomitant heart rate changes in rabbits (Oryctolagus cuniculus). *Behavioral Neuroscience*, *115*(5), 1029–38.

Pullum, G. K., & Scholz, B. (2010). Recursion and the infinitude claim. *Recursion in Human Language*, 103–104.

Qin, L., Chimoto, S., Sakai, M., Wang, J., & Sato, Y. (2007). Comparison between offset and onset responses of primary auditory cortex ON-OFF neurons in awake cats. *Journal of Neurophysiology*, *97*(5), 3421–31. doi:10.1152/jn.00184.2007

Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience*, *13*(6), 766–85. doi:10.1162/08989290152541430

Raghavachari, S., Kahana, M. J., Rizzuto, D. S., Caplan, J. B., Kirschen, M. P., Bourgeois, B., … Lisman, J. E. (2001). Gating of human theta oscillations by a working memory task. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(9), 3175–83.

Raij, T., McEvoy, L., Mäkelä, J. P., & Hari, R. (1997). Human auditory cortex is activated by omissions of auditory stimuli. *Brain Research*, *745*(1-2), 134–43.

Rainer, G., Rao, S. C., & Miller, E. K. (1999). Prospective coding for objects in primate prefrontal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(13), 5493–505.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive field effects. *Nature Neuroscience*, *2*(1), 79–87.

Rao, R. P., & Sejnowski, T. J. (2001). Spike-timing-dependent Hebbian plasticity as temporal difference learning. *Neural Computation*, *13*(10), 2221–37. doi:10.1162/089976601750541787

Rao, R., & Sejnowski, T. (2002). Predictive Coding, Cortical Feedback, and Spike-Timing Dependent Plasticity. In R. P. N. Rao, B. A. Olshausen, & M. S. Lewicki (Eds.), *Probabilistic models of the brain: perception and neural function* (MIT Press.). Cambridge.

Raymond, J., Shapiro, K., & Arnell, K. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology. Human Perception and Performance*, *18*(3), 849–60.

Reale, R. A., & Imig, T. J. (1980). Tonotopic organization in auditory cortex of the cat. *The Journal of Comparative Neurology*, *192*(2), 265–91. doi:10.1002/cne.901920207

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863.

Remillard, G. (2008). Implicit learning of second- , third- , and fourth-order adjacent and nonadjacent sequential dependencies. *Experimental Psychology*, *61*(3), 400 – 424. doi:10.1080/17470210701210999

Remillard, G. (2010). Implicit learning of fifth- and sixth-order sequential probabilities. *Memory & Cognition*, *38*(7), 905–15. doi:10.3758/MC.38.7.905

Remillard, G., & Clark, J. M. (2001). Implicit learning of first-, second-, and third-order transition probabilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 483–498.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: current research and theory* (New York: ., pp. 64–99).

Riggall, A. C., & Postle, B. R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(38), 12990–8. doi:10.1523/JNEUROSCI.1892-12.2012

Rinne, T., Antila, S., & Winkler, I. (2001). Mismatch negativity is unaffected by top-down predictive information. *Neuroreport*, *12*(10), 2209–13.

Ritter, W., Sussman, E., Deacon, D., Cowan, N., & Vaughan, H. G. (1999a). Two cognitive systems simultaneously prepared for opposite events. *Psychophysiology*, *36*(6), 835–8.

Ritter, W., Sussman, E., Deacon, D., Cowan, N., & Vaughan, H. G. (1999b). Two cognitive systems simultaneously prepared for opposite events. *Psychophysiology*, *36*(6), 835–8.

Rogers, J., & Pullum, G. K. (2011). Aural Pattern Recognition Experiments and the Subregular Hierarchy. *Journal of Logic, Language and Information*, *20*(3), 329–342. doi:10.1007/s10849-011-9140-2

Rohrmeier, M., Fu, Q., & Dienes, Z. (2012). Implicit learning of recursive context-free grammars. *PloS One*, *7*(10), e45885. doi:10.1371/journal.pone.0045885

Rose, S., & Walker, R. (2003). A Typology of Consonant Agreement as Correspondence Sharon Rose and Rachel Walker University of California, San Diego and University of Southern California July 2003, (July).

Rosenblath, W. (1899). Uber einen bemerkenswerten Fall von Hirnerschutterung. *Dutch Arch Klein Med*, *64*, 406–420.

Rothschild, G., Nelken, I., & Mizrahi, A. (2010). Functional organization and population dynamics in the mouse primary auditory cortex. *Nature Neuroscience*, *13*(3), 353–360. doi:10.1038/nn.2484

Roy, J. E., Riesenhuber, M., Poggio, T. A., & Miller, E. K. (2010). Prefrontal cortex activity during flexible categorization. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(25), 8519–28. doi:10.1523/JNEUROSCI.4837-09.2010

Rünger, D., & Frensch, P. A. (2008). How incidental sequence learning creates reportable knowledge: the role of unexpected events. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(5), 1011–26. doi:10.1037/a0012942

Sabri, M., & Campbell, K. B. (2001). Effects of sequential and temporal probability of deviant occurrence on mismatch negativity. *Cognitive Brain Research*, *12*, 171–180.

Saffran, J. R., Aslin, R., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Listening (and Leaming) out of the Comer of Your Ear. *Psychological Science*, *8*(2), 101–106.

Salisbury, D., Squires, N., Ibel, S., & Maloney, T. (1992). Auditory event-related potentials during stage 2 NREM sleep in humans. *Journal of Sleep Research*, *1*(4), 251–257. doi:10.1111/j.1365-2869.1992.tb00047.x

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology*, *62*(6), 437–448.

Sato, Y., Yabe, H., Hiruma, T., Sutoh, T., Shinozaki, N., Tadayoshi, N., & Kaneko, S. (2000). The effect of deviant stimulus probability on the human mismatch process. *Neuroreport*, *11*(7), 3703–3708.

Sawaguchi, T., & Goldman-Rakic, P. S. (1991). D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science (New York, N.Y.)*, *251*(4996), 947–50.

Sawaguchi, T., & Goldman-Rakic, P. S. (1994). The role of D1-dopamine receptor in working memory: local injections of dopamine antagonists into the prefrontal cortex of rhesus monkeys performing an oculomotor delayed-response task. *J Neurophysiol*, *71*(2), 515–528.

Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation for action in the parietal cortex of the monkey. *Nature*, *415*(6874), 918–22. doi:10.1038/415918a

Schreiner, C. E., & Mendelson, J. R. (1990). Functional topography of cat primary auditory cortex: distribution of integrated excitation. *Journal of Neurophysiology*, *64*(5), 1442–59.

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behavioral and Brain Functions : BBF*, *6*, 24. doi:10.1186/1744-9081-6-24

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, *275*(5306), 1593–9.

Schvaneveldt, R. W., & Chase, W. G. (1969). Sequential effects in choice reaction time. *Journal of Experimental Psychology*, *80*(1), 1–8. doi:10.1037/h0027144

Sculthorpe, L. D., Ouellet, D. R., & Campbell, K. B. (2009). MMN elicitation during natural sleep to violations of an auditory pattern. *Brain Research*, *1290*, 52–62. doi:10.1016/j.brainres.2009.06.013

Sederberg, P., Kahana, M., Howard, M., Donner, E. J., & Madsen, J. R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. *The Journal of Neuroscience*, *23*(34), 10809 –10814.

Seidenberg, M. S. (1997). Language Acquisition and Use: Learning and Applying Probabilistic Constraints. *Science*, *275*(5306), 1599–1603. doi:10.1126/science.275.5306.1599

Sekar, K., Findley, W. M., Poeppel, D., & Llinas, R. R. (2013). Cortical response tracking the conscious experience of threshold duration visual stimuli indicates visual perception is all or none. *Proceedings of the National Academy of Sciences*, *110*(14). doi:10.1073/pnas.1302229110

Self, M. W., Kooijmans, R. N., Super, H., Lamme, V. a., & Roelfsema, P. R. (2012). Different glutamate receptors convey feedforward and recurrent processing in macaque V1. *Proceedings of the National Academy of Sciences*, 1–6. doi:10.1073/pnas.1119527109

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *8*(10), 1391–400. doi:10.1038/nn1549

Servan-Schreiber, D., Carter, C., & Bruno, R. (1998). Dopamine and the mechanisms of cognition: Part II. D-amphetamine effects in human subjects performing a selective attention task. *Biological Psychiatry*.

Seth, A. (2007). Models of consciousness. *Scholarpedia*, *2*(1), 1328.

Shadlen, M. N. (2011). Consciousness as a decision to engage. In S. Dehaene & Y. Christen (Eds.), *Characterizing Consciousness: From Cognition to the Clinic? Research and Perspectives in Neurosciences*. (pp. 27–46). Berlin: Springer Verlag.

Shieber, S. M. (1985). Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, *8*(3), 333–343. doi:10.1007/BF00630917

Shima, K., Isoda, M., Mushiake, H., & Tanji, J. (2006). Categorization of behavioural sequences in the prefrontal cortex. *Nature*, *445*(7125), 315–318. doi:10.1038/nature05470

Shinozaki, N., Yabe, H., Sutoh, T., Hiruma, T., & Kaneko, S. (1998). Somatosensory automatic responses to deviant stimuli. *Cognitive Brain Research*, *7*(2), 165–71.

Shipp, S., Adams, R. a, & Friston, K. J. (2013a). Reflections on agranular architecture: predictive coding in the motor cortex. *Trends in Neurosciences*, *36*(12), 706–16. doi:10.1016/j.tins.2013.09.004

Shipp, S., Adams, R. a., & Friston, K. J. (2013b). Reflections on agranular architecture: predictive coding in the motor cortex. *Trends in Neurosciences*, 1–11. doi:10.1016/j.tins.2013.09.004

Siegel, J. J., Kalmbach, B., Chitwood, R. a, & Mauk, M. D. (2012). Persistent activity in a cortical-to-subcortical circuit: bridging the temporal gap in trace eyelid conditioning. *Journal of Neurophysiology*, *107*(1), 50–64. doi:10.1152/jn.00689.2011

Siegel, M., Warden, M. R., & Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences*, *106*(50), 21341–6. doi:10.1073/pnas.0908193106

Sigala, N., Kusunoki, M., Nimmo-Smith, I., Gaffan, D., & Duncan, J. S. (2008). Hierarchical coding for sequential task events in the monkey prefrontal cortex. *Proceedings of the National Academy of Sciences*, *105*(33), 11969–74. doi:10.1073/pnas.0802569105

Simpson, T. P., Manara, a R., Kane, N. M., Barton, R. L., Rowlands, C. a, & Butler, S. R. (2002). Effect of propofol anaesthesia on the event-related potential mismatch negativity and the auditory-evoked potential N1. *British Journal of Anaesthesia*, *89*(3), 382–8.

Singh, S., Barto, A. G., & Chentanez, N. (2004). Intrinsically motivated reinforcement learning. *18th Annual Conference on Neural Information Processing Systems (NIPS)*.

Singh, S., Lewis, R., & Barto, A. (2009). Where do rewards come from. *Proceedings of the Annual Conference of the Cognitive Science Society*, 2601–2606.

Sokolov, E. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology*, *25*, 545–580.

Solomon, P. R., Stowe, G. T., & Pendlbeury, W. W. (1989). Disrupted eyelid conditioning in a patient with damage to cerebellar afferents. *Behavioral Neuroscience*, *103*(4), 898–902.

Sporns, O., Tononi, G., & Edelman, G. M. (2000). Connectivity and complexity: the relationship between neuroanatomy and brain dynamics. *Neural Networks : The Official Journal of the International Neural Network Society*, *13*(8-9), 909–22.

Spratling, M. W. (2010). Predictive Coding as a Model of Response Properties in Cortical Area V1. *The Journal of Neuroscience*, *30*(9), 3531–3543. doi:10.1523/JNEUROSCI.4911-09.2010

Spratling, M. W. (2012). Predictive coding as a model of the V1 saliency map hypothesis. *Neural Networks : The Official Journal of the International Neural Network Society*, *26*, 7–28. doi:10.1016/j.neunet.2011.10.002

Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, *193*(4258), 1142–1146.

Squires, N., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387–401.

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. S. (2013). Dynamic Coding for Cognitive Control in Prefrontal Cortex. *Neuron*, *78*(2), 364–375. doi:10.1016/j.neuron.2013.01.039

Striem-Amit, E., Cohen, L., Dehaene, S., & Amedi, A. (2012). Reading with sounds: sensory substitution selectively activates the visual word form area in the blind. *Neuron*, *76*(3), 640–52. doi:10.1016/j.neuron.2012.08.026

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*(9), 1004–1006. doi:10.1038/nn.2163

Sussman, E., Ritter, W., & Vaughan, H. G. (1998). Predictability of stimulus deviance and the mismatch negativity. *Neuroreport*, *9*(18), 4167–70.

Sussman, E., Winkler, I., & Schröger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin & Review*, *10*(3), 630–7.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, *3*(1), 9–44. doi:10.1007/BF00115009

Sutton, R. S., & Barto, a. G. (1998a). Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks*, *9*(5), 1054–1054. doi:10.1109/TNN.1998.712192

Sutton, R. S., & Barto, A. G. (1998b). *Reinforcement learning: An introduction* (p. 360). Cambridge Univ Press. doi:10.1016/S1364-6613(99)01331-5

Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, *112*(1-2), 181–211. doi:10.1016/S0004-3702(99)00052-1

Sutton, S., Braren, M., Zubin, J., & John, E. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187.

Sysoeva, O., Takegata, R., & Näätänen, R. (2006). Pre-attentive representation of sound duration in the human brain. *Psychophysiology*, *43*(3), 272–6. doi:10.1111/j.1469-8986.2006.00397.x

Taine, H. (1882). *De l'intelligence* (Paris: Hac.).

Tales, A., Newton, P., Troscianko, T., & Butler, S. (1999). Mismatch negativity in the visual modality. *NeuroReport*, *10*(16), 3363–3367.

Tanji, J., Shima, K., & Mushiake, H. (2007). Concept-based behavioral planning and the lateral prefrontal cortex. *Trends in Cognitive Sciences*, *11*(12), 528–534. doi:10.1016/j.tics.2007.09.007

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science (New York, N.Y.)*, *332*(6033), 1054–9. doi:10.1126/science.1196404

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: statistics, structure, and abstraction. *Science (New York, N.Y.)*, *331*(6022), 1279–85. doi:10.1126/science.1192788

Thomson, A. M., & Lamy, C. (2007). Functional maps of neocortical local circuitry. *Frontiers in Neuroscience*, *1*(1), 19–42. doi:10.3389/neuro.01.1.1.002.2007

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies.* (Macmillan.).

Tikhonravov, D., Neuvonen, T., Pertovaara, a, Savioja, K., Ruusuvirta, T., Näätänen, R., & Carlson, S. (2010). Dose-related effects of memantine on a mismatch negativity-like response in anesthetized rats. *Neuroscience*, *167*(4), 1175–82. doi:10.1016/j.neuroscience.2010.03.014

Tikhonravov, D., Neuvonen, T., Pertovaara, A., Savioja, K., Ruusuvirta, T., Näätänen, R., & Carlson, S. (2008). Effects of an NMDA-receptor antagonist MK-801 on an MMN-like response recorded in anesthetized rats. *Brain Research*, *1203*(Haartmaninkatu 8), 97–102. doi:10.1016/j.brainres.2008.02.006

Tishby, N., Pereira, F., & Bialek, W. (2000). The information bottleneck method. *arXiv Preprint physics/0004057*, 1–16.

Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior Expectation Mediates Neural Adaptation to Repeated Sounds in the Auditory Cortex: An MEG Study. *Journal of Neuroscience*, *31*(25), 9118–9123. doi:10.1523/JNEUROSCI.1425-11.2011

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208.

Topka, H., Valls-Solé, J., Massaquoi, S. G., & Hallett, M. (1993). Deficit in classical conditioning in patients with cerebellar degeneration. *Brain : A Journal of Neurology*, *116 (Pt 4*, 961–9.

Treisman, M. (1963). Temporal discrimination and the indifference interval. Implications for a model of the "internal clock". *Psychological Monographs*, *77*(13), 1–31.

Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *The Journal of Neuroscience*, *24*(46), 10440–53. doi:10.1523/JNEUROSCI.1905-04.2004

Van Aalderen-Smeets, S. I., Oostenveld, R., & Schwarzbach, J. (2006). Investigating neurophysiological correlates of metacontrast masking with magnetoencephalography. *Advances in Cognitive Psychology*, *2*(1), 21–35. doi:10.2478/v10053-008-0042-z

Van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(22). doi:10.1073/pnas.1117465109

Van Gaal, S., & Lamme, V. A. F. (2012). Unconscious high-level information processing: implication for neurobiological theories of consciousness. *The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *18*(3), 287–301. doi:10.1177/1073858411404079

Van Opstal, F., Van Laeken, N., Verguts, T., Van Dijck, J.-P., De Vos, F., Goethals, I., & Fias, W. (2014). Correlation between individual differences in striatal dopamine and in visual consciousness. *Current Biology*, 1–16.

Van Schouwenburg, M. R., den Ouden, H. E. M., & Cools, R. (2010). The human basal ganglia modulate frontal-posterior connectivity during attention shifting. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(29), 9910–8. doi:10.1523/JNEUROSCI.1111-10.2010

Varela, J., Sen, K., Gibson, J. R., Fost, J., Abbott, L. F., & Nelson, S. B. (1997). A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *The Journal of Neuroscience*, *17*(20), 7926–40.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(6), 1656–74.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(April), 748–751.

Vogel, E. K., Mccollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(November), 500–503. doi:10.1038/nature04171

Volkov, I. O., & Galazjuk, a V. (1991). Formation of spike response to sound tones in cat auditory cortex neurons: interaction of excitatory and inhibitory effects. *Neuroscience*, *43*(2-3), 307–21.

Wacongne, C., Labyt, E., Van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences*, *108*(51), 20754–59. doi:10.1073/pnas.1117807108

Walter, W. ., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. *Nature*, *203*(4943), 380–384.

Wang, X.-J. (1999). Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(21), 9587–603.

Wang, X.-J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neurosciences*, *24*(8), 455–63.

Warden, M. R., & Miller, E. K. (2007). The representation of multiple objects in prefrontal neuronal delay activity. *Cerebral Cortex (New York, N.Y. : 1991)*, *17 Suppl 1*, i41–50. doi:10.1093/cercor/bhm070

Weible, a P., Weiss, C., & Disterhoft, J. F. (2007). Connections of the caudal anterior cingulate cortex in rabbit: neural circuitry participating in the acquisition of trace eyeblink conditioning. *Neuroscience*, *145*(1), 288–302. doi:10.1016/j.neuroscience.2006.11.046

Weible, A. P., McEchron, M. D., & Disterhoft, J. F. (2000). Cortical involvement in acquisition and extinction of trace eyeblink conditioning. *Behavioral Neuroscience*, *114*(6), 1058–67.

Weiss, C., Bouwmeester, H., Power, J. M., & Disterhoft, J. F. (1999). Hippocampal lesions prevent trace eyeblink conditioning in the freely moving rat. *Behavioural Brain Research*, *99*(2), 123–32.

Willingham, D. B. (1999). Implicit motor sequence learning is not purely perceptual. *Memory & Cognition*, *27*(3), 561–72.

Winkler, I. (2007). Interpreting the mismatch negativity (MMN). *Journal of Psychophysiology*, *21*(3-4), 147–163.

Winkler, I., Cowan, N., Csépe, V., Czigler, I., & Näätänen, R. (1996). Interactions between transient and long-term auditory memory as reflected by the mismatch negativity. *Journal of Cognitive Neuroscience*, *8*(5), 403–415.

Winkler, I., Takegata, R., & Sussman, E. (2005). Event-related brain potentials reveal multiple stages in the perceptual organization of sound. *Brain Research. Cognitive Brain Research*, *25*(1), 291–9. doi:10.1016/j.cogbrainres.2005.06.005

Woodruff-Pak, D. S., & Disterhoft, J. F. (2008). Where is the trace in trace conditioning? *Trends in Neurosciences*, *31*(2), 105–12. doi:10.1016/j.tins.2007.11.006

Woodruff-Pak, D. S., Lavond, D. G., & Thompson, R. F. (1985). Trace conditioning: abolished by cerebellar nuclear lesions but not lateral cerebellar cortex aspirations. *Brain Research*, *348*(2), 249–60.

Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *NeuroReport*, *8*, 1971–1974.

Zeki, S. (2003). The disunity of consciousness. *Trends in Cognitive Sciences*, *7*(5), 214–218. doi:10.1016/S1364-6613(03)00081-0