ELECTROPHYSIOLOGICAL MEASURES OF RESIDUAL LANGUAGE COMPREHENSION TO IMPROVE DIAGNOSTIC ACCURACY IN PROLONGED DISORDERS OF CONSCIOUSNESS

By

CONSUELO VIDAL GRAN

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Centre for Human Brain Health (CHBH)

School of Psychology

College of Life and Environmental Sciences

University of Birmingham

September 2020

Abstract

Behavioural scales are the clinical standard to assess consciousness in patients diagnosed with disorders of consciousness (DOC), although patients remain unresponsive to environmental stimuli beyond reflexes. Electrophysiological studies have identified modulation of brain activity in some patients when following commands; however, command-following requires complex brain processing that most patients lack due to the severity of their brain injury. Passive EEG paradigms differentiating automatic from strategic processing have detected residual cognitive processing in DOC patients. The local-global paradigm measures trial by trial violations of local expectations (short time-scales), which elicit a mismatch negativity (MMN) response that reflects automatic processing; whereas violations of global expectations (longer time-scales), that are generated using contextual information, show a positive ERP component (P3b) that accounts for controlled processing. Following the same rationale, this thesis includes four experiments that investigate the influence of semantic local and global expectations on word processing, aiming to assess residual language comprehension in DOC patients. By employing a relatedness proportion paradigm in a semantic priming task, we provide a 'local' context within trials (i.e., related/unrelated word-pairs), as we simultaneously manipulate global expectations by providing a context across the task (i.e., prime validity: cueing participants about the probability of a related target following the prime). We first report three behavioural visual studies in healthy participants, which suggest that individuals show greater priming effects in high validity contexts relative to low validity contexts (i.e., relatedness proportion effects), as individuals use the global context strategically to generate expectations about

the target. Strategic involvement is supported by self-report measures, as only individuals that applied conscious strategies while performing the task showed these behavioural global effects. In a subsequent study we investigate the neural correlates of the generation of local and global expectations in this paradigm. The results show an earlier 'local' prediction error ERP effect around 250ms; followed by a later 'global' effect approximately at 350ms that interacted with the global context. We then adapt the task for DOC patients to identify auditory markers of strategic expectancy generation in this paradigm. The results in the healthy group suggest an early 'local' ERP effect (around 350ms), showing a prediction error ERP signal; and a later effect (around 550ms) reflecting violations of the 'global' semantic context i.e., larger error signal for unexpected targets in high validity relative to low validity context. Moreover, we present two DOC patient cases, where we detect a 'local' ERP effect in a patient (MCS diagnosis), whereas no ERP effect is observed in the other case (VS/UWS diagnosis). In conclusion, we propose this auditory ERP paradigm as a tool to detect residual language comprehension in DOC patients, as it detects hierarchical differentiated effects for strategic and non-strategic semantic expectations, similar to the local-global paradigm but in the clinically relevant domain of language processing.

Dedication

To Andrés

For your infinite support

I’m grateful we embarked on this risky adventure

that became a home.

Acknowledgements

I would like to thank all those who have contributed in any way to the completion of this thesis and supported me throughout the process of completing my PhD. First, I would like to express my infinite gratitude to my two genius supervisors Dr. Damian Cruse and Professor Howard Bowman, for their constant support and advice, and most importantly that they always trusted in my abilities and challenged me to keep learning and improving. I would also like to thank Dr. Katrien Segaert and Dr. Tristan Bekinschtein for their valuable feedback while examining the present thesis.

I wish to thank all the members of the Disorders of Consciousness Lab; Davinia, Damian, Leah, Rodi, Sean, Sara, Davide, Roya, Mel, Ruth, as I feel so grateful to be part of the best research lab where I could have done my PhD. Thank you for your fabulous friendship, and permanent academic and personal support; thanks for all the laughs and our adventures at the office and outside of it. Moreover, I would like to express my gratitude to everyone at the School of Psychology and the CHBH building, who contributed directly or indirectly to both the research project and my experience at the University, special thanks to everyone at the 111 CHBH office, Seb, Tom, Al, Ben, and so many more people that made my PhD experience so special.

I would like to show my gratitude to a direct contributor of this project, “Becas Chile, ANID” from the Government of Chile, which funded my studies and supported me financially during the PhD. I am grateful for the opportunity to develop my career in a prestigious University, and to have the personal experience of living abroad.

I would like to give special thanks to my parents Marta and Pedro Pablo, my sister María José, my niece Emma, my brother-in-law Francisco (Pollo) and my husband Andrés, who are my emotional support, and all my achievements are thanks to them. Thanks to my lovely grandparents who will always be present, Nona, Gigi, Titi, Tata. Furthermore, to all my friends that have supported me during my whole PhD here in the UK, special thanks to Yoyo, Fonchi, Sara, Ceci, Leander, Nico, S; and from Chile: Ichi, Challo, Barbarita, Ale, Tere and all the w. promedios. I would also like to show my gratitude to Carlos M (KS) for his constant support during the process of writing this thesis.

List of contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of tables

Chapter 2

Chapter 3

Chapter 4

CHAPTER 1: GENERAL INTRODUCTION

The clinical challenge

The term acquired brain injury (ABI) includes any type of brain injury (Turner-Strokes, 2003) that causes structural damage to the brain, is acquired after birth, and has either a traumatic or non-traumatic aetiology (Griffiths, 2006). Brain injuries caused by head trauma are defined as traumatic brain injury (TBI) (National Institute for Health and Care Excellence, 2014) and can be classified into two types; first, closed injuries where the primary internal damage is generated by a traumatic event (e.g., head impact in a road traffic accident), followed by internal secondary effects caused by the trauma, including anoxia, haemorrhage or swelling (Griffiths, 2006). Secondly, penetrating injuries in which the head is impacted by external objects creating an open injury (Griffiths, 2006).

On the contrary, there are several types of brain injuries of non-traumatic aetiology that are caused by internal accidents such as cerebrovascular events (i.e., stroke or subarachnoid haemorrhage); hypoxic brain damage which refers to an oxygen deficiency in the brain; metabolic or toxic disorders; or other infections in the brain such as encephalitis (Turner-Strokes, 2003; Griffiths, 2006). In any type of ABI, individuals are susceptible to develop cognitive deficits as a consequence of the damage caused by the injury; and the scope of the damage, which may be localised (e.g., stroke) or diffuse (e.g., trauma secondary effects), may influence the level of cognitive deficit (Turner-Strokes, 2003). Moreover, the severity of a brain injury can be classified as mild, moderate, or severe, which is assessed when patients arrive to emergency centres (National Institute for Health and Care Excellence, 2014); here, the severity of the brain injury is usually evaluated with the Glasgow Coma Scale, GCS (Teasdale & Jennett, 1974). The GCS measures patient's responsiveness to the environment including three domains: eye opening (spontaneous, in response to speech and painful stimulation); motor response (in response to commands and painful stimulation); and verbal response to speech (Teasdale & Jennett, 1974). The GCS has a minimum of 3 points and a maximum of 15 points, patients must score above 12 to be classified as mild brain injury, where they are able to consciously respond to commands and painful stimulation, and they maintain communication abilities. Moreover, moderate brain injury requires a score between 9 and 12, in which patients respond to verbal commands, are able to withdraw from painful stimulation, and show confused/incoherent communication. For severe brain injury, the Glasgow score should be below 8, where patients are usually unresponsive to verbal and noxious stimulation, with no eye opening and they do not show verbal responses (National Institute for Health and Care Excellence, 2014; Teasdale & Jennett, 1974). Moreover, a computerized tomography (CT) scan is part of the diagnosis, which is used to assess the level of structural damage caused by the brain injury (National Institute for Health and Care Excellence, 2014).

Within the UK, TBI and stroke are the most common causes of acquired brain injury (Turner-Strokes, 2003). The incidence rate in England for patients with head injury admitted to an emergency department was reported to be 229.4 per 100,000 from 2001 to 2002, and 229.1 per 100,000 in 2002 to 2003 (Tennant, 2005). Another study reported that 453 per 100,000 patients each year attended to emergency services as estimated from a range of 6 years (1997-2003); moreover, the rate of patients presenting moderate to severe head injury in the UK was around 40 per 100,000, which corresponds to 10.9% of all cases of head injury (Yates, Williams, Harris, Round, & Jenkins, 2006). More recently, the National Institute for Health and Care Excellence (2014) stated that the yearly estimate attendance to emergency services due to head injury in the UK is around 1.4 million patients, and around 200,000 are hospitalized for this reason. Regarding TBI, a recent study estimated that globally 69 million people experience TBI each year (Dewan et al., 2018). Furthermore, another global estimation of TBI yielded an annually proportion of 295 per 100,000 considering all ages, where the incidence varied by the severity of the TBI indicating 224 per 100,000 for mild; 23 per 100,000 for moderate; and 13 per 100,000 for severe TBI (Nguyen et al., 2016). Patients suffering moderate to severe brain injury are at risk to experience an impaired state of consciousness, such as a comatose state (Blume, Del Giudice, Wislowska, Lechinger, & Schabus, 2015; Turner-Strokes, 2003).

Impaired consciousness

In order to understand the meaning of an impaired or altered state of consciousness, the concept of consciousness must be addressed. First, there is no consensus on a unified definition of consciousness within the scientific community (Owen, 2008), as consciousness is an elusive concept that can be approached from diverse disciplines, such as philosophy, psychology, neuroscience, and many others (De Sousa, 2013; Seth, 2010; Zeman, 2001). Each discipline investigates consciousness from their own perspective, such as its origins, how it is explained with the laws of the universe or how do we explain it from a biological point of view (Koch, Massimini, Boly, & Tononi, 2016). Therefore, achieving an integrated definition of consciousness across disciplines has been described as a major scientific challenge (Zeman, 2001). Despite these limitations in its conceptualization, some general definitions have been proposed; for example, consciousness have been described as the ability to be aware of one's own mental states and to be able to communicate them to other individuals (Gazzaniga, Ivry & Mangun, 2009). Moreover, being conscious involves having an experience and others can assume a person is conscious when they are behaviourally awake, meaning that they show voluntary behaviour and are capable of self-reporting their experience (Koch et al., 2016). Zeman (2001) explains the concept of consciousness from three different perspectives. First, consciousness can be understood as a continuum that ranges from waking states, then sleep to unconscious states; and consciousness may be lost, depressed, or regained. Second, consciousness described as the content of subjective experience which occurs at any given time. Third, consciousness defined as mind, which refers to mental states with propositional content, such as hopes, beliefs, expectations, etc. (Zeman, 2001). Similarly, consciousness has been described as a phenomenon that is manifested in a specific level within a continuum (e.g., comatose state) and with specific contents of personal conscious experiences (Seth, 2010; De Sousa, 2013; Cavanna, Shah, Eddy, Williams, & Rickards, 2011).

Research on altered states of consciousness seeks to understand what would be broadly agreed as both conscious and unconscious states (Owen, 2008). Since achieving a general definition of consciousness is challenging, research on altered states of consciousness generally addresses the definition of consciousness from a clinical perspective. Thus, consciousness is considered to be composed of wakefulness, which is manifested through arousal; and awareness of the environment and the self (Blume et al., 2015; Cavanna et al., 2011; Laureys, Boly, Moonen, & Maquet, 2009; Zeman, 2001; Posner & Plum, 2007). Both aspects of consciousness have distinctive features, where the level of wakefulness (i.e., arousal) refers to general brain activity that indicates how responsive an individual is to external stimuli (Goldfine & Schiff, 2011; Cavanna et al., 2011); and awareness refers to a complex process that integrates the activity of simultaneous brain networks, in order to process external stimuli that are received by the different sensory modalities, such as visual, auditory, etc. (Goldfine & Schiff, 2011). Therefore, a fully wakeful state would be defined as an individual who shows signs of volition, processes external stimuli, and has an intact level of arousal (Goldfine & Schiff, 2011). Evidence that stems from neuroimaging and lesion studies identifies some candidates for the neural correlates of consciousness (NCC), that can be either content-specific or full consciousness (Koch et al., 2016). From a neural perspective, neurons that shape the dorsal tegmentum of the midbrain and the pons are thought to be involved in the waking state, and their trajectories include the central thalamus and basal forebrain, which then send information via thalamo-cortical and basal forebrain-cortical projections (Goldfine & Schiff, 2011). Moreover, parts of the brainstem, thalamus and postero-medial cortex have been proposed as support areas for maintaining a conscious state; however, there is no definitive answer to which brain networks could be responsible for the manifestation of consciousness (Koch et al., 2016).

As mentioned above, one consequence of severe brain injury is that patients may result in a state of coma, which can be described as an acute period of unconsciousness

(Blume et al., 2015), where individuals are not able to be aroused, thus are behaviourally unresponsive; maintain their eyes closed; and lack any evidence of being aware of themselves or the environment (Cavanna et al., 2011; Young, 2009; Owen, 2008; Zeman, 2001). As individuals in coma cannot be aroused, they are not able to behaviourally respond to external stimulation (Owen, 2008). Furthermore, patients generally show withdrawal responses when presented with noxious stimuli; however, these responses may not be present in deep comatose states (Posner, Plum, Saper, & Schiff, 2007). After a period of two to four weeks, patients in coma commonly start recovering consciousness, whereas others may not survive (Owen, 2008; Laureys, Owen & Schiff, 2004). From the group of patients that recover after the comatose state, some will regain their normal cognitive functions; while others might transit into a different state of impaired consciousness, where they continue to be behaviourally unresponsive to the environment (Blume et al., 2015; Owen, 2008).

Disorders of consciousness

Jennet and Plum (1972) referred to the need to name and describe the state that follows a comatose state, where patients are no longer unconscious but have a profound alteration in consciousness. The authors named this state as Persistent Vegetative State (VS). The alternative name Unresponsive Wakefulness Syndrome (UWS) was suggested (Laureys et al., 2010) and nowadays both terms are used in research and clinical contexts, including the acronym of both terms combined: VS/UWS; which will be used in the present thesis, as both terms refer to the same diagnostic label. VS/UWS is a clinical syndrome where patients evidence wakefulness, but there are no behavioural signs of awareness (Zeman, 2001). When patients progress from a comatose state to a VS/UWS, they recover their sleep wake cycles (Blume et al., 2015; Zeman, 2001) and are able to open and close their eyes (Owen, 2008). When observing VS/UWS patients, they seem to be awake and they respond to painful stimuli; nevertheless, they appear to be behaviourally unaware of themselves and/or the environment and are not able to communicate with others or evidence any sign of purposeful action (Royal College of Physicians, 2020; Blume et al., 2015; Cavanna et al., 2011; Zeman, 2001). The observer (e.g., clinician, family member) may be confounded by the patient's eye opening and reflexive responses, which can be incorrectly interpreted as voluntary behaviour (Owen, 2008).

The VS/UWS syndrome can be transitory or irreversible (Blume et al., 2015; Owen, 2008), and it can be classified as *persistent*, which means that the patient had remained in that state for more than one month after the severe brain injury event; or it can be classified as *permanent*, lasting more than three months for non-traumatic brain injury and a minimum of twelve months for traumatic brain injury aetiologies (Owen, 2008; Multi-Society Task Force, 1994). However, the most recent clinical guidelines for prolonged disorders of consciousness in the UK, proposes the terms *continuing VS/UWS* for patients meeting the criteria for VS/UWS for more than 4 weeks; and *chronic VS/UWS* for patients diagnosed as VS/UWS for more than 3 months for non-traumatic aetiologies, and more than 12 months for traumatic aetiologies (Royal College of Physicians, 2020). Furthermore, *permanent VS/UWS* can be diagnosed when a patient has received the diagnosis of *chronic VS/UWS* for more than 6 months and has not shown any change or improvement; in addition, the permanent VS/UWS diagnosis can only be made by a qualified prolonged DOC Physician (Royal College of Physicians, 2020). To date, there is no specific method to predict the likelihood of recovery or the permanence in a VS/UWS (Royal College of Physicians, 2020; Zeman, 2001).

Some patients do not meet the criteria for being diagnosed as VS/UWS, but still maintain an impaired state of consciousness; therefore, Giacino and colleagues (2002) proposed to create the classification Minimally Conscious State (MCS). The main difference between patients in VS/UWS and MCS is that the latter show behavioural evidence of consciousness yet are not able to reproduce it consistently over time (Giacino et al., 2002). The MCS can be defined as "a condition of severely altered consciousness in which minimal but definite behavioural evidence of self or environmental awareness is demonstrated" (Giacino et al., 2002, p. 350-351). Some patients in VS/UWS may go through MCS before recovering, whereas others may indefinitely remain in an MCS (Owen, 2008). An indicator of recovery from MCS corresponds to the ability of communicating and/or using objects in a functional manner (Blume et al., 2015; Owen, 2008). In the same way as VS/UWS, the diagnosis of MCS has been recently adapted to *continuing MCS*, where patients must show signs of inconsistent interaction with their environment for more than 4 weeks; for *chronic MCS*, the MCS diagnosis has continued for more than 9 months for non-traumatic, and more than 18 months for traumatic aetiologies. Permanent MCS diagnosis can only be carried out by an expert physician and the patient must have mantained the chronic diagnosis for more than 6 months with no changes or improvements (Royal College of Physicians, 2020).

There are no updated estimates about the incidence and prevalence of DOC in the UK, mainly because there is high variability in the diagnostic criteria (Van Erp et al., 2014; Beaumont & Kenealy, 2005). For example, a prevalence review study by Van Erp and colleagues (2014) found methodological differences between the different studies that were analysed, where there was not a unified diagnostic criterion. Only 5 of 14 studies included the MCS label as a separate diagnosis from VS/UWS; therefore, it is not possible to obtain a reliable prevalence rate. However, it was estimated that there were around 4,000 – 16,000 patients in nursing homes in the UK with prolonged DOC and the number increases three times for MCS patients (Houses of Parliament, 2015). The numbers of DOC patients are growing as medical technology advances, and because of these advances fewer patients die as a result of a brain injury (Blume et al., 2015; Monti, Laureys, & Owen, 2010).

The complete clinical assessment includes a review of the clinical history of the patient to determine the cause of the brain injury, and a differential diagnosis to discard other conditions. Moreover, the assesment also contains a review of the patient's medication, a detailed neurological assessment (i.e., reflexes, responses to noxious stimuli, etc.), and standard brain imaging (i.e., CT or MRI) when deemed necessary (Royal College of Physicians, 2020). After the standard medical evaluation has been carried out, clinicians conduct the clinical observation of behavioural responses (Royal College of Physicians, 2020; Owen & Coleman, 2008).

Diagnosing consciousness

For making a clinical behavioural diagnosis, clinicians must grade the severity of the impairment of consciousness by establishing whether the patient can respond to external stimuli by performing overt responses (Owen & Coleman, 2008). For this purpose, the most commonly used diagnostic behavioural scales are the Glasgow Coma Score (GCS) for coma, and the JFK Coma Recovery Scale (CRS-R) for VS/UWS and MCS (Teasdale & Jennett, 1974; Kalmar & Giacino, 2005). As indicated by the Royal College of Physicians (2020), other validated scales that are also used in the UK to diagnose prolonged disorders of consciousness are the Wessex Head Injury Matrix (WHIM; Shiel et al., 2000) and Sensory Modality Assessment and Rehabilitation Technique (SMART; Gill-Thwaites & Munday, 2004). In general, standardized behavioural scales are designed to measure auditory, visual, verbal and motor abilities; where clinicians instruct individuals to follow specific commands and assign a score according to the individual's performance (Koch et al., 2016). The GCS, specifically, measures the depth of the altered state of consciousness and assesses three behavioural domains: eye opening, verbal responses, and motor responses to stimulation (Teasdale & Jennett, 1974). In each domain the clinician has to rate the patient's response to stimulation using a range that goes from normal responses to no response. For example, in the eye-opening domain, the clinician would rate whether the patient presents spontaneous eye opening (normal response); eye opening in response to speech; in response to pain; or no eye opening at all. The same type of assessment is conducted in the other two domains (Teasdale & Jennett, 1974). The CRS-R has diagnostic, prognostic, and treatment planning features. The scale is composed by 25 items ordered in 6 subscales (auditory, visual, motor, oromotor, communication, and arousal) that assess whether the patients can perform certain actions, such as object recognition, orientation, and movement to command, among others (Kalmar & Giacino, 2005). Similar to GCS, clinicians must rate the patient's behavioural responses using a range that goes from present to absent responses (Kalmar & Giacino, 2005).

The clinical evaluation of patients' behaviour involves not only the use of behavioural scales, but also information collected from different sources, such as clinical notes, observation from family members, health workers and trained physicians that conduct the assessments (Royal College of Physicians, 2020). However, the formal behavioural scales have a fundamental role when establishing the specific diagnosis, and these should be repeated over time in order to make a diagnostic decision (Royal College of Physicians, 2020). Specifically, for the diagnosis of VS/UWS the patient must show a lack of behavioural responses regarding awareness, language processing, and voluntary behaviours. For example, absence of behavioural responses would mean that the individual is not responding verbally to the clinician, not being able to voluntary manipulate objects, and not being orientated in time and space, etc. (Royal College of Physicians, 2013). For the MCS diagnosis the patient must evidence responses that imply awareness of the self or the environment. These responses should be replicated in time to be accounted for the diagnosis and should be at least one of the following behaviours: respond to simple verbal commands from the clinician, gestural or verbal yes/no responses, verbally respond, and/or any voluntary behaviour (Giacino et al., 2002).

Behavioural assessment limitations

Behavioural measures depend on patient's behaviour, thus their own capacity to show their level of awareness (Blume et al., 2015). However, patients undergo high levels of physical disability or motor impairment, so it is unlikely that they can show signs of awareness through their overt behaviour (Owen, 2008). This contradiction has caused some patients to be misdiagnosed as vegetative state (Andrews, Murphy, Munday, & Littlewood, 1996). Several authors had investigated the rate of misdiagnosis in DOC patients. Andrews and colleagues (1996) investigated a group of 40 VS/UWS patients, who received occupational therapy sessions for 6 weeks. The intervention aimed at assessing the patients' ability to respond to sensory stimulation and commands. The response was assessed through a buzzer button press or direct gaze towards an object. Patients were taught to respond with yes (one press) and no (two presses) and were asked biographical questions that were later contrasted with family responses. The authors found that 17 out of 40 (43%) patients were misdiagnosed as VS/UWS, as they were able to follow commands using the button press (Andrews et al., 1996). Childs, Mercer, and Childs (1993) had previously reported a similar rate of misdiagnosis, where 18 out of 49 patients (37%) were diagnosed inaccurately. More recently, it was reported that 41% of patients were misdiagnosed as VS/UWS with behavioural assessments using CRS-R (Schnakers et al., 2009). This study included a total of 103 patients, whose diagnoses were determined through repeated clinical behavioural observations made by a clinical team, where 44 patients were diagnosed as VS/UWS; 41 as MCS; and 18 with uncertain diagnosis. Of the patients in a VS/UWS, 18 (41%) showed signs of awareness when assessed with CRS-R.

Voluntary eye movements were considered to be the main sign implying conscious awareness. In the group of patients with undefined diagnosis, 16 (89%) showed signs of awareness; therefore, the authors suggested that patients with uncertain diagnosis are more likely to be in MCS, rather than VS/UWS (Schnakers et al., 2009). Even though these previous studies showed relatively similar percentages of patients misdiagnosed as VS/UWS, i.e., 43% (Andrews et al., 1996), 37% (Childs et al., 1993), 41% (Schnakers et al., 2009); only the latter was conducted after the MCS diagnostic label existed in the clinical practice; thus, as the authors included both classifications in their analysis (VS/UWS and MCS) their estimation becomes a more accurate and updated analysis of the misdiagnosis rate within DOC.

The high rate of misdiagnosis of DOC in clinical settings indicates the need to identify other sources of information in addition to behavioural measures (Monti et al., 2010); as some patients may be unable to perform overt motor responses to external stimuli due to motor impairment, rendering behavioural assessments insufficient to ensure an accurate clinical diagnosis (Cruse et al., 2012; Cruse & Owen, 2010, Coleman et al., 2007). The need for accuracy in the clinical diagnosis of patients with DOC relies on several factors (Graham et al., 2015; Faugeras et al., 2012). First, having greater accuracy regarding the presence or absence of awareness in DOC patients facilitates medical management, by clarifying which patients are more likely to benefit from rehabilitation programs (Faugeras et al., 2012). Higher accuracy in this matter would also be advantageous from a familial and economical point of view (Racine, Rodrigue, Bernat, Riopelle, & Shemie, 2010). From the patient's family perspective, an accurate diagnosis could reduce the uncertainty regarding the patient's state and thus contribute to reduce the psychological distress that relatives experience (Fins, 2013). Furthermore, from an economical perspective, the allocated resources in health care from governments could be directed to patients that would benefit from rehabilitation; on the opposite, more accuracy would help clarify the cases where end-of-life decisions should be debated (Racine et al., 2010). From the patient's perspective, the detection of covert awareness and the possibility of communication would contribute to knowing whether patients experience pain, how they experience their disability and knowing about their psychological wellbeing (Graham et al., 2015).

Neuroimaging for improved diagnostic accuracy

Some patients have been found to retain residual preserved awareness, even though they score as unresponsive on behavioural measures. Research with functional neuroimaging have identified residual cognitive functions in some patients diagnosed as VS/UWS or MCS when they are requested to follow commands, even though these abilities had not been revealed through standard behavioural approaches. A pivotal study with functional magnetic resonance imaging (fMRI) by Owen and colleagues (2006) used two imagery tasks to assess the ability to follow commands in a patient diagnosed as VS/UWS. The patient received verbal instructions to imagine herself, first playing tennis and then navigating through her own house. Both tasks were also performed by healthy individuals (N: 12) and the results showed similar patterns of hemodynamic activation when comparing this group with the patient. The results showed activation in the same brain areas for both healthy individuals and the patient, such as activity in the supplementary motor area (SMA) in the tennis task, and activation in parahippocampal gyrus (PPA), posterior parietal-lobe (PPC), and lateral premotor cortex (PMC) for the navigation task (Owen et al., 2006).

The imagery tasks allow assessing the individual´s responses without requiring any behavioural motor response, which is appropriate for DOC patients since they have motor impairment, thus are unable to respond overtly (Owen, 2008). The ability to follow commands from verbal instructions require individuals to deploy a range of cognitive abilities, such as language comprehension, attention, working memory resources, response selection, etc. (Fernández-Espejo & Owen, 2013; Cruse, et al., 2011). Therefore, following commands and being able to respond through the modulation of brain activity corresponds to a strong indicator that an individual is consciously aware and responding to the task intentionally (Owen et al., 2006).

The imagery tasks were then replicated by Monti and colleagues (2010), where a group of patients (5 out of 54 – 9.2%) showed modulation of their brain activity when performing imagery tasks. One of these patients was even able to use both the tennis and navigation tasks to give yes or no responses. One of the tasks served as a yes response and the other as a no response; therefore, the patient was able to establish communication with the researchers as they asked questions that required yes/no responses (Monti et al., 2010). These results indicate that some patients might respond to commands by modulating their brain activity, which would imply that they preserve awareness even though they were behaviourally diagnosed as unaware. Hence, these techniques have the potential to provide an objective measure of awareness in clinical settings for patients diagnosed with DOC (Fernández-Espejo & Owen, 2013; Monti et al., 2010; Cruse & Owen, 2010).

Although functional magnetic resonance imaging (fMRI) is a promising complement to behavioural measures for the diagnosis of DOC, it represents a difficult challenge (Laureys & Schiff, 2012; Cruse, Monti, & Owen, 2011; Cruse et al., 2011;). Not all patients are able to access fMRI because it is expensive; requires the patient to be transported to the scanning facility, which can produce physical stress for the patient; and cannot be used with patients who have metal plates or pins (Peterson, Cruse, Naci, Weijer, & Owen, 2015; Cruse et al., 2011). On the contrary, Electroencephalography (EEG) represents both an alternative and a complement to fMRI as a testing tool for the presence of awareness in patients with DOC (Sitt et al., 2014; Cruse & Owen, 2010). EEG is an electrophysiological method that uses electrodes around the scalp to measure voltage fluctuations in the brain; and these variations correspond to signals that are direct reflections of neural oscillations in the cortex (Cohen, 2014). Compared to fMRI, EEG is more applicable in clinical settings because of its portability, reduced costs, all patients are suitable for its application, and it can be used at the bedside (Cruse & Owen, 2010; Cruse et al., 2011). A study by Cruse and colleagues (2011) recorded EEG responses to imagery tasks that required command following in VS/UWS patients and healthy controls. Individuals were requested to imagine squeezing their right hand and then relaxing it or wiggling their toes and then relaxing them. From a total of 16 patients diagnosed as VS/UWS, 3 (19%) showed EEG responses that evidenced command following, suggesting that these patients were capable of modulating their brain activity when they were requested to imagine either squeezing their hand or wiggling their toes (Cruse et al., 2011).

The tasks described above attempt to assess command following through imagery tasks in patients with DOC using EEG or fMRI (Cruse et al., 2011; Monti et al., 2010; Owen et al., 2006), which are of high complexity for this group of patients, as several cognitive functions are required to be involved while performing the task. For example, to follow an instruction an individual needs to comprehend what is being asked, pay attention to the instructions, and sustain attention throughout the task; moreover, a simultaneous and successful use of these abilities altogether indicates conscious awareness (Cruse et al., 2011). However, DOC patients have severe brain injury so it is likely that some of their cognitive functions would be disrupted due to the damage caused to the axonal fibres and brain structures (Griffiths, 2006). Imagery tasks are considered as active paradigms, where the effort relies on the patient's ability to follow commands and imagine what is requested, which may be too complex for this group of patients (Cruse et al., 2011). Alternatively, some paradigms are passive, meaning that individuals are not required to follow commands, as the stimuli are designed and presented to specifically measure certain residual cognitive abilities. For example, Perrin and colleagues (2006) designed a paradigm where series of names and tones were auditorily presented to DOC patients, and the patient's own name was randomly included within the sequence. The task did not require patients to actively follow command, as they were only expected to passive listening to the stimuli. The authors found a differentiated EEG response to the patient's own name with respect to other names.

Conscious vs nonconscious brain processes

To assess whether a patient with DOC still preserves traces of cognitive processing or awareness, it is essential to design paradigms that allow discriminating automatic nonconscious brain processes, from cognitive processes that require conscious processing (Faugeras et al., 2012); moreover, using passive and active tasks in combination has been proposed as an efficient method to test this differentiation (Blume et al., 2015; Schnakers et al., 2008). A key paradigm in the field of disorders of consciousness is the Local-Global paradigm (Bekinschtein et al., 2009), which differentiates between automatic and conscious processing, besides being active and also can be used passively (King, Gramfort, Schurger, Naccache, & Dehaene, 2014). This EEG oddball paradigm aims at measuring differentiated responses between violations of local expectations (within each trial) and violations of global expectations that are given by the context in which stimuli are presented (blocks across the task). For this purpose, the paradigm uses a series of 5 consecutive tones in each trial and measures the EEG responses to the fifth tone. In *local-standard trials (LS)* the fifth tone has identical pitch to the previous four, whereas the fifth tone in *local-deviant trials (LD)* has a different pitch from previous tones. Each trial represents the local context, which unfolds in a short time-scale and the detection of deviant trials elicits an automatic brain response that does not rely on conscious processing (Bekinschtein et al., 2009). On a longer time-scale, series of both LD and LS trials are presented in different proportions (either 80% or 20%) within a block to establish the global context. Therefore, local trials (deviant or standard) may become either *global-deviant (GD)* or *global-standard (GS)* according to the context in which they are presented. For example, a LD trial presented within a block where

80% of trials are also LD becomes a global-standard as the majority of items are LD, even though on a local level this trial is deviant. Consequently, a locally standard (LS) trial within the same block will be simultaneously a global deviant (GD) trial, as it corresponds to a violation of that block's global regularity because the majority of trials are LD. Evidence suggests that the detection of global-deviant (GD) trials relies on conscious processing, since individuals must be aware of the block regularity to detect the global violations of expectations (Bekinschtein et al., 2009), this evidence is reviewed below.

In the local-global paradigm, event-related potentials (ERPs) are used to analyse responses to auditory stimuli (Bekinschtein et al., 2009). ERPs are a widely used technique in EEG research and it refers to electrophysiological responses time-locked to a specific stimulus (Luck, 2014). The ERP response is a direct measurement of neural activity, so ERP responses that occur under different experimental conditions are contrasted to observe the signal differences that are given by these experimental manipulations (Luck, 2014). Researchers usually classify their results into ERP components, which describe the waveform that may have negative or positive deflections in voltage; components also specify the time (milliseconds) in which the difference between conditions reaches the peak. For example, N400 would indicate a negative going waveform that peaks around 400 milliseconds (Luck, 2014; Kutas and Federmeier, 2011).

In healthy participants, the local-global paradigm elicits an initial Mismatch Negativity (MMN) in response to local violations of expectations, followed by a centro-parietal positivity around 300ms post-stimulus - usually referred as the P3b component – that is a result of global violations of expectation (Faugeras et al. 2012; King et al., 2014; El

Karoui et al., 2015; Bekinschtein et al., 2009). The MMN ERP component peaks around 150-250ms and has been associated with changes in auditory stimuli that does not rely on conscious awareness (Näätänen, Pakarinen, Rinne, & Takegata, 2004). In the original local-global manipulation, healthy participants performed either a counting task, mind wandering or a distracting task. All individuals showed the MMN response to local violations of expectations; however, the P3b to global violations of expectations was only evident in the group of participants that were instructed to count the global-deviant trials across the task. The P3b response dropped for individuals instructed to mind-wander and decreased even more for participants that conducted a visual distraction task. The P3b response therefore only occurs in this paradigm when individuals are aware of the global context; hence, it accounts as a useful electrophysiological marker of conscious access (Bekinschtein et al., 2009). Moreover, the authors tested the paradigm in DOC patients (8 patients: 4 VS/UWS, 4 MCS) where 3 out of 4 VS/UWS patients showed the local effect (MMN), whilst there was no indicator of global effect (P3b). Regarding MCS patients a local effect was observed in all individuals, whereas 3 out of 4 patients showed a significant global effect (P3b); who all three regained consciousness after a few weeks (Bekinschtein et al., 2009).

The global effect as a marker of conscious processing

The idea that conscious processing is necessary for the presence of a global effect, has been supported by studies testing individuals under conditions that are not considered as conscious states, such as anaesthesia and sleep, where an individual's receptiveness to environmental stimuli ceases momentarily (Strauss et al., 2015). A study by Nourski and colleagues (2018) tested the local-global paradigm on individuals during anaesthesia. Specifically, the authors used intracranial recordings in epileptic patients that underwent surgery and presented the task in three distinct moments: first, before surgery when patients were awake; next, while patients were still awake and were being sedated (using propofol); and then after sedation when patients lost consciousness and remained unresponsive. The stimuli were composed by vowel letters (a and i), instead of pitch changes of auditory tones as the original manipulation (Bekinschtein et al., 2009). The electrodes were placed in auditory cortical areas and the prefrontal cortex (PFC). The authors reported both local and global ERP effects (MMN followed by a P300) in awake patients in all regions of interest (ROI). Patients during sedation (responsive) showed the local effect in the auditory areas but not in the PFC, whereas there was no evidence for the global effect in any ROI. Regarding anesthetised (unresponsive) patients, the local effect was only present in certain areas of the auditory cortex and the global affect was not detected (Nourski et al., 2018). Another study by Strauss and colleagues (2015) investigated violations of local and global auditory regularities in sleep. Their participants also performed the vowel local-global paradigm in several conditions: before sleep (wake-pre), during sleep (N1, N2 and REM stages) and after sleep (wake-post). The ERP results for the awake conditions (pre and post) were consistent with previous studies as an MMN response was elicited for violations of local regularities, whilst a P3b response emerged in response to violations of global regularities. Regarding the sleep conditions (N1, N2 and REM), the results provided evidence for a local effect – MMN response – in all sleep conditions; however, there was no detection of a global effect in any of the sleep conditions (Strauss et al., 2015).

Both studies investigating local and global effects during sleep or anaesthesia that were mentioned above, provide evidence for the MMN as a marker of unconscious processing, where the brain continues to process stimuli even though the individual remains unresponsive. Moreover, both studies also provide evidence for the P300 as a marker of conscious processing, as it is detected when individuals remain responsive and then is no longer detected when individuals enter an unresponsive state (Nourski et al., 2018; Strauss et al., 2015; Chennu & Bekinschtein, 2012). Although we can use sleep and anaesthesia to study brain responses in unconscious states, direct comparisons with DOC should be avoided as they correspond to distinctive phenomena; specifically, DOC are a result of extensive brain damage and it is not entirely clear whether they remain in a state of full unconsciousness or still preserve residual cognitive abilities (Chennu & Bekinschtein, 2012). Nevertheless, the evidence indicates that detecting global violations requires consciousness.

Although the global-effect is considered a strong marker of conscious access and has repeatedly provided evidence on its high specificity for detecting conscious processing in DOC patients; the global effect has shown low sensitivity as not all patients that are demonstrably conscious have presented this response (Faugeras et al., 2012). For example, a study that tested the local-global paradigm in both healthy controls and DOC patients, even though they reported global effects in all healthy participants using single-subject analyses, the effect was not detected in all patients that were diagnosed as conscious (Faugeras et al., 2012). Specifically, on a single subject level, 100% of healthy controls (8/8), 53.8% (7/13) of conscious patients, 14.3% (4/28) of MCS, and 8% (2/24) of VS/UWS patients evidenced a P3b as a result of the global effect (Faugeras et al., 2012). The global effect is predominantly present in healthy individuals that are attending to the task by counting or monitoring the global deviant trials across an entire block (Bekinschtein et al., 2009), which represents a complex task for DOC patients that have regular fluctuations of their attentional resources as a consequence of their brain injury (Faugeras et al., 2012).

Residual language comprehension as a measure of consciousness

Several authors have mentioned the notion of having a multidomain ERP assessment to evaluate a range of cognitive abilities in the same manner as it is done in other clinical populations (Rohaut et al., 2015; Faugeras et al., 2012). Using tasks that differentiate between automatic and conscious strategic processing of stimuli (e.g. local-global paradigm), in combination with the assessment of other cognitive processes, such as semantic processing, could be a useful tool for diagnosis of DOC (Faugeras et al., 2012).

Strategic cognitive control refers to the ability to direct the attentional resources in a flexible manner according to task demands to reach a determined goal (Dymowski, Owens, Ponsford, & Willmott, 2015; van Gaal, De Lange, & Cohen, 2012; Horga & Maia, 2012). For example, in a task's performance the detection of errors can account for strategic processing (van Gaal et al., 2012). Moreover, Seth, Dienes, Cleeremans, Overgaard, and Pessoa (2008) argued that studies that aim at measuring the presence of awareness should combine brain measures that can be explained by underlying behavioural measures, which may include objective measures; subjective measures; and strategic control. Conscious behaviour could be expressed through these three types of behavioural measures in experimental tasks: *objective measures* when individuals make decisions based on experimental conditions, for example being able to predict upcoming stimuli based on task demands; *strategic control* when individuals make use of the information provided in the task's instructions for their performance advantage; and *subjective measures* when individuals directly report their experiences (Seth et al., 2008).

In addition, paradigms that do not rely on patients fully following commands (i.e. imagery tasks) should be considered when developing future tasks, because it may result in a clearer estimate of the residual cognition in DOC patients (Cruse et al., 2014). Moreover, knowing to what extent the patient retains residual language processing may be useful for the patient´s prognosis and could contribute as a clinical marker of the subsequent return of awareness (Coleman et al., 2007). Language expression and comprehension are considered as main aspects of awareness, because knowing that someone is aware is primarily determined by their ability of communication (Owen & Coleman, 2008). For example, behavioural scales that are used for the clinical diagnosis of DOC are presented to the patient verbally by the clinician; therefore, their performance would not only rely on their ability to execute the command, but also on their ability to understand the verbal instructions, and understanding speech is in itself a conscious experience (Teasdale & Jennett, 1974; Kalmar & Giacino, 2005).

The value of assessments of language comprehension in DOC is evident from studies that have investigated residual language abilities in DOC patients using both fMRI and EEG methods (Cruse et al., 2014). A study using fMRI by Coleman and colleagues (2007) intended to detect residual language processes that support speech comprehension. The authors aimed to distinguish three levels of language processing: *auditory* (auditory stimuli vs silent baseline), *perceptual* (clear vs incomprehensible speech), and *semantic* (ambiguous vs unambiguous sentences). The ROI determined for each condition were the following: the superior temporal plane was identified for the auditory condition; the superior temporal sulcus and left inferior frontal gyrus for the perceptual condition; and at the higher level in the semantic condition, it was identified the left posterior inferior temporal areas and the left inferior frontal gyrus. From the DOC patients' group (VS/UWS, MCS, MSC emergence), 2 patients that emerged from an MCS showed auditory and perceptual responses, but only one of them showed significant activation in the temporal lobe for the semantic condition. Next, 5 out of 12 VS/UWS or MCS showed temporal lobe activation for both the auditory and perceptual levels, and 3 patients showed activity in the temporal lobe for the semantic level. From these results, the authors proposed that DOC patients may preserve what they had named as "islands of preserved cognitive function" (Coleman et al., 2007, p. 2495).

These islands could be useful for the diagnosis, as it is not possible to observe them through the patient's behaviour (Coleman et al., 2007). The same paradigm was later tested on a larger sample of patients with both VS/UWS and MCS diagnosis. The results detected significant temporal lobe activity in 19 of 41 patients (including VS/UWS and MCS) in the perceptual condition; moreover, at the top level of the semantic hierarchy (semantic condition) 4 patients showed signs of semantic processing, although only in some ROI, which the authors attribute to the changes in brain structure due to the brain injury, obstructing a direct comparison with healthy individuals (Coleman et al., 2009).

Following the same rationale, another study by Beukema and colleagues (2016) investigated a similar hierarchical semantic paradigm adapted to measure ERP responses, however, the authors used a normative-association priming task (implemented by Cruse et al., 2014) for the semantic condition instead of sentences (the semantic priming effect is further explained in the next section). The prime-target word-pairs were preceded by two unintelligible words and participants were instructed to mentally decide whether the word-pair was related or unrelated. At the perceptual level, 7 of 16 patients showed significant effects (single-subject analysis), which correspond to a greater negative ERP signal for words relative to noise, as indicated by the healthy participants group analyses. At the highest hierarchical level (semantic condition), a significant effect was detected in one patient, which was a larger ERP signal for unrelated word-pairs with respect to related word-pairs, as it was revealed by healthy participants group analyses (Beukema et al., 2016). This effect is comparable to the classic N400 effect that is elicited when contrasting related and unrelated words-pairs (Kutas & Federmeier, 2011). Although this paradigm measures hierarchical language processing that goes from hearing sounds, then distinguish between noise and words, and showing differentiated responses for related and unrelated word-pairs (Beukema et al., 2016); the higher hierarchical level (i.e., semantic priming) has no support for being a direct measurement of conscious processing, and evidence for this statement is provided in the next chapter and throughout this thesis.

Expectations as markers of conscious processes

In the same way as expectancy mechanisms influence the processing of auditory regularities, several authors have proposed that language comprehension is influenced by expectations at multiple hierarchical levels (Lewis & Bastiaansen, 2015; Ylinen et al., 2016; Lau, Holcomb, & Kuperberg, 2013; Hutchison, 2007; Kuperberg & Jaeger, 2016). The idea that the brain actively generates expectations about the information it encounters stems from the predictive coding theory, which posits that the brain functions in a Bayesian manner, by contrasting prior expectations with sensory input (Clark, 2013). One of the main postulates of this theory is that the brain actively processes information by generating expectations about upcoming stimuli; and these expectations are based on previous knowledge of the world and the context where the information unfolds (Heilbron & Chait, 2018; Clark, 2013). Hence, throughout our lives the brain is thought to build models of the world containing the predictable features of the environment (Friston, 2010); as the steady stream of incoming sensory information develops, these models are constantly being contrasted with the sensory input. When the models detect the unpredictable components of the stimuli, an error signal is generated in the brain (Huang & Rao, 2011). The higher levels in the hierarchical organization of the brain generate inferences about the upcoming sensory input by contrasting the model and current context via top-down connections. Otherwise, when the sensory input is unpredicted by the model, an error signal is transmitted from lower to higher levels of the hierarchy via bottom-up connections (Heilbron & Chait, 2018; Friston & Kiebel, 2009; Rao & Ballard, 1999), forcing the model to adapt to the perceived information, and thus acting as a prior hypothesis for the inference of the subsequent stimuli (Clark, 2013). As a result, the new learned information is progressively incorporated into the models to reduce the amount of prediction error, in order to achieve successful perception (Friston, 2010; Clark, 2013). Expectations are considered to occur at several locations in the cortex, in different levels of the hierarchy and simultaneously or at different time scales across the cortical layers (Clark, 2013).

The local-global paradigm is a clear example of how expectations influence behaviour and neural processes in a bidirectional bottom-up and top-down manner (Bekinschtein et al., 2009). On a local-level, the perception of a local-deviant stimulus elicits a prediction error response (bottom-up processing) as it violates the expected regularity, and on a neural level the error is manifested as an MMN, where the ERP amplitude is larger for unexpected stimuli with respect to expected stimuli. On a global level, perceiving a global-deviant trial also generates a prediction error signal (P3b ERP response); however, as global trials change according to the context in which they are presented, individuals are required to make conscious inferences about global trials; therefore, the prediction errors stem from the higher levels of the hierarchy (top-down processing) (Heilbron & Chait, 2018; Friston & Kiebel, 2009; Rao & Ballard, 1999).

Expectations in language comprehension

Similarly, language comprehension is thought to be influenced by inference mechanisms, that have a role in facilitating language processing (Kuperberg & Jaeger, 2016). The semantic priming paradigm has been broadly used in language research to investigate how expectancy mechanisms influence language comprehension (Hutchison et al., 2013; Kuperberg, Jaeger, 2016). Semantic priming tasks use prime-target word-pairs that can be either semantically related or unrelated; where the Semantic Priming Effect (SPE) refers to individuals showing faster behavioural responses to targets (e.g., DOG) that are preceded by related primes (e.g., CAT) than unrelated primes (e.g., LAMP) (Hutchison et al., 2013; Gulan & Valerjev, 2010).

There are contrasting views on whether semantic priming is generated by automatic or expectancy mechanisms (Lau et al., 2013; Hutchison, 2007). The theory that supports the automatic point of view is the spreading activation processing, which was first proposed by Posner, Snyder, and Solso (1975). This theory explains that when an individual is presented with a prime, there is an automatic activation in the brain of the semantic representation for that word; then, this information automatically spreads to pathways of associated words, so when the target is presented, there is an identification of the word. Therefore, this facilitation in the identification of the target would explain why individuals are faster to respond to related targets, whereas in the unrelated targets the response is slower since there is no identification produced (Hutchison, 2007; Yap, Hutchison, & Tan, 2016). In contrast, the semantic priming effect has been explained as a result of an expectancy process (Hutchison, 2007; Keefe & Neely, 1990). This theory explains that the individual generates a set of possible words that would be related to the prime. When the target is presented, if it matches with the possibilities, it will generate facilitation; whereas if it does not match it will produce inhibition (Hutchison, 2007). When a target is related to the prime an individual would show a faster reaction time as they are able to make a prediction of the upcoming stimuli, while if the word is unrelated, this would produce a delay in the response as no prediction is possible (Kuperberg & Jaeger, 2016). The semantic priming effect can be measured using tasks such as lexical decision tasks, naming tasks, phrasal decision tasks and speech monitoring (Kuperberg & Jaeger, 2016). A more extensive explanation about automatic versus expectancy mechanisms, and the type of tasks used in semantic priming paradigms is detailed in Chapter 2.

Semantic Priming experiments using ERPs have found a negative deflection in voltage peaking around 400ms, which is broadly known as the N400 component (Kutas & Hillyard, 1980; Kutas & Federmeier, 2011). The N400 effect refers to a reduction in the ERP amplitude for related targets in contrast to unrelated targets (Kutas & Federmeier, 2011). Under a predictive coding perspective, the N400 would reflect the prediction error response as unrelated targets (unexpected) show a greater amplitude with respect to related targets (expected) (Fitz & Chang, 2019; Rabovsky & McRae, 2014).

Language processing in the context of disorders of consciousness

In the context of disorders of consciousness some authors had investigated the semantic priming paradigm using ERPs (Rohaut et al., 2015; Cruse et al., 2014). A study by Cruse and colleagues (2014) investigated the N400 effect elicited by a semantic priming task in a group of healthy participants. In the first experiment individuals either completed an overt condition (make a related or unrelated response on each trial); a covert condition (think about their response); or a passive condition (simply attend to individual words). The results showed that the overt condition elicited greater N400 effects on a group level with respect to the covert and passive conditions. On a single-subject level the N400 detection varied depending on the participants' level of involvement in the task, as the N400 effect was detected in 75% (overt), 58% (covert), and 0% (passive) of participants. The lack of sensitivity for the passive condition in healthy participants on a single-subject level, poses a difficulty to assess language processing with the N400 effect in DOC patients, as they are not able to provide overt responses. Thus, the authors conducted a second experiment where they used normative associative prime-target word pairs (Cruse et al., 2014). Unlike semantically related word-pairs where prime and target are usually from the same semantic category (e.g., dog - fox) (Keefe & Neely, 1990), associative priming is constructed by presenting individuals with a prime word and asking them to write the first associated word that they can think of, so then on a group level, the target becomes the most common word across participants (e.g. 80% of participants mentioning the word "dog" when seeing the word "cat") (Hutchison et al., 2013). For example, in the Semantic Priming Project (Hutchison et al., 2013) a database was created using associated prime-target word-pairs (taken from Nelson, McEvoy, & Schreiber, 1999), where these were validated in a sample of 768 individuals (more details in chapter 2).

Using normative associated data, Cruse and colleagues (2014) conducted the semantic priming task on healthy participants with only the passive condition. On a single subject-level, the N400 effect was detected in 50% of individuals. When comparing both experiments, using normative associated word-pairs as opposed to semantically related word-pairs increased the sensitivity to detect the N400 on a single-subject level from 0% to 50%; therefore, the authors proposed that a way of increasing the sensitivity to detect N400 effects in DOC patients (single-subjects) is using normative associative prime-target word-pairs (Cruse et al., 2014). As previously mentioned, this normatively associated stimuli were included as the highest level of a semantic hierarchy in a group of DOC patients, where the N400 effect was only observed in one MCS patient; thus, the authors concluded that detecting N400 effects in DOC patients on a single-subject level lacks clinical utility due to its low sensitivity in detecting the effect (Beukema et al., 2016).

Other studies have reported N400 effects in some DOC patients (Rohaut et al., 2015; Kotchoubey et al., 2005; Schoenle & Witzke, 2004; Steppacher et al., 2013). A study by Rohaut and colleagues (2015) used a classical auditory semantic priming task where individuals passively listened to the words. In healthy participants, the results detected an N400 effect followed by a late positive complex (LPC) – known as P600 (Kutas & Federmeier, 2011) -, both effects showed greater voltage values for incongruent (i.e., unrelated) targets in contrast to congruent (i.e., related) targets. These effects were also source localised to the right temporal pole and middle frontal gyrus for the N400 effect; and the LPC was localised to the inferior frontal gyrus and right fusiform gyrus. On a group level, an N400 effect was found in the DOC, VS/UWS and MCS groups; but there was no significant LPC observed in any of the three groups. On a single-subject level, 42.1% of healthy participants presented an N400 effect, and 31.6% a significant LPC. Regarding patients, a significant N400 effect was detected in 20.7% of all DOC patients (6/29; 5 MCS and 1 VS/UWS). Whereas a significant LPC was observed in the same percentage of patients (6/29; 5 MCS and 1 VS/UWS), and two other MCS patients showed an LPC using a regression analysis. As the LPC was only detected in controls and MCS patients, relative to VS/UWS patients, where only one patient showed the effect, the authors propose that the LPC reflects semantic representation. Therefore, the authors suggested a hierarchical processing of words where the N400 would account for the lower level (non-conscious) while the LPC would represent a higher level (conscious) that shares similar characteristics to the role of the P3b in the local-global paradigm (Rohaut et al., 2015).

Supporting the view that N400 effects can be observed in individuals that remain in an unconscious state, a study by Rämä and colleagues (2010) tested an auditory semantic priming paradigm in 13 comatose patients, where one group had temporal cortex damage and the other group had an intact temporal cortex. The results yielded significant N400 effects for the intact temporal areas, in contrast with the group having damage in those areas, where no N400 effects were detected. These results provide evidence for the involvement of temporal areas in speech processing and support the idea of the N400 being a result of automatic processing; as coma corresponds to the most severe level of altered states of consciousness, where only automatic processing occurs and complex conscious processing is deemed unlikely (Rämä et al., 2010). The semantic priming task therefore does not represent a reliable measure for seeking conscious residual language processing in DOC patients, as this task does not allow differentiating between automatic and conscious processing, and this differentiation has been pointed as crucial for the design of tasks assessing residual cognitive functions in DOC patients (Faugeras et al, 2012).

Some studies have provided evidence for the presence of N400 effects under states that are not entirely conscious, such as sleep. For example, a study investigating semantic processing during sleep reported N400 effects when individuals were asleep and heard related and unrelated words-pairs. The ERP signal was greater for unrelated words relative to related words, and this effect was comparable to the effect detected in awake participants (Perrin, Bastuji & Garcia-Larrea, 2002). On the contrary, a study by Kallionpää and colleagues (2018) found N400 effects in awake individuals that were presented with sentences that had congruous or incongruous ending. The incongruous words elicited a greater ERP signal with respect to congruous endings. The same participants performed the task under anaesthesia, where they received either Dexmedetomidine or Propofol. The results showed that once individuals were sedated and unresponsive, no N400 effect was detected (Kallionpää et al., 2018). These results support the view that N400 effects are only present in fully conscious participants, contrary to studies endorsing the view that the N400 effect can be present in unconscious states, such as coma (Rämä et al., 2010) or sleep (Perrin, Bastuji & Garcia-Larrea, 2002). Other studies have provided evidence for N400 effects when awake healthy participants are not aware of prime presentation, by employing masked semantic priming manipulations (Rolke, Heil, Streb, & Hennighausen, 2001; Kiefer

& Brendel, 2006; Kiefer, 2002; Deacon, Hewitt, Yang, & Nagata, 2000). Masked priming refers to an experimental manipulation where the primes are presented for a brief period of time in between a pattern mask (e.g., string of letters before and after the prime); where individuals usually report not being aware of the presence of the prime (Kiefer, 2002). A study by Kiefer (2002) reported the detection of N400 effects in both masked and unmasked conditions of a semantic priming task, where the target was presented 67ms after the prime onset (SOA). These results suggest that N400 effects can be a result of automatic processing for two reasons: First, priming effects at such short SOAs are thought to rely on automatic processing due to spreading activation (Hill, Strube, Roesch-Ely, & Weisbrod, 2002). Secondly, individuals are not aware of the primes in the masked condition, therefore, the elicited N400 effect is induced by a prime that lacks conscious access, providing evidence that the concept of semantic priming is not sufficient to assess signs of conscious processing (Kiefer, 2002).

The N400 literature shows evidence in favour of both conscious and nonconscious/automatic processing involvement; therefore, by not having clarity regarding the N400 nature, no inferences can be made on whether an individual has conscious processing, just by the mere presence of N400 effects. The ERP N400 effects followed by an LPC (P600) that are elicited by a semantic priming task were proposed as two effects reflecting hierarchical semantic processing; where the N400 represents nonconscious/automatic processing of words and the LPC accounts for conscious processing of semantic content (Rohaut & Naccache, 2017). This profile was proposed in accordance to the MMN that is elicited by violations of local regularities, reflecting nonconscious/automatic processing and the P3b that is detected for violations of global regularities, which has been defined as a marker of conscious processing; as was proposed in the local-global paradigm (Rohaut & Naccache, 2017; Bekinschtein et al., 2009). However, when comparing the local-global task with a semantic priming task, the latter only represents a local level of processing. Unrelated targets would represent local errors, as they violate any expectation that may have been produced from the prime, while a related target would fulfil the expectation generated from the prime; both in the same manner as a local trial would be either deviant or standard. A semantic priming task lacks the global manipulation that the local-global paradigm implements, where the proportion of locally deviant and standard trials are manipulated across blocks, allowing individuals to generate conscious expectations given by the global context in which stimuli is presented (Bekinschtein et al., 2009).

One approach to separating automatic/non-strategic from strategic processing in language comprehension is the relatedness proportion paradigm (RP) in a semantic priming task, which manipulates the overall proportion of semantically associated (related) prime-targets pairs across the experiment. This manipulation aims at generating a global context across the experiment, so individuals become aware when they are in a high prime validity context, in contrast with a low prime validity context (Hutchison, 2007; Keefe & Neely, 1990; Lau et al., 2013). If the proportion of related pairs increases across the task (e.g. 80% related and 20% unrelated), the size of priming effects will thus increase (Hutchison, 2007), so as the magnitude of N400 effects (Lau et al., 2013). As the relatedness proportion facilitates semantic priming in high validity contexts, relative to low validity contexts; individuals can use this context to know the likelihood of the target being related or unrelated to the prime. In other words, individuals use the prime to predict the target based on the context (Hutchison, 2007; Keefe & Neely, 1990).

From an expectation point of view, the local-global paradigm and the relatedness proportion paradigm in a semantic priming task operate under the same rationale. In the local-global paradigm, local expectations (within trial) can be either standard or deviant, so as the targets in the RP paradigm that can be either related (standard), or unrelated (deviant). Moreover, on a more complex level, global expectations (across blocks) in the local-global manipulation will depend on the proportion assigned to each type of trial within the block, in the same way as global expectations in the RP task, where the prime validity context will define which type of trial is the most and least expected, see figure 1.1. Therefore, we propose that the relatedness proportion paradigm (RP) in a semantic priming task will allow testing separately local semantic expectations (non-strategic) from global semantic expectations, which are given by the context in which stimuli are presented, thus, requiring the use of strategic and conscious processing. Therefore, prediction of upcoming semantic input could be a potential candidate to measure residual conscious language comprehension in patients diagnosed with disorders of consciousness. The following chapter will outline in detail the arguments for the utility of the relatedness proportion paradigm.

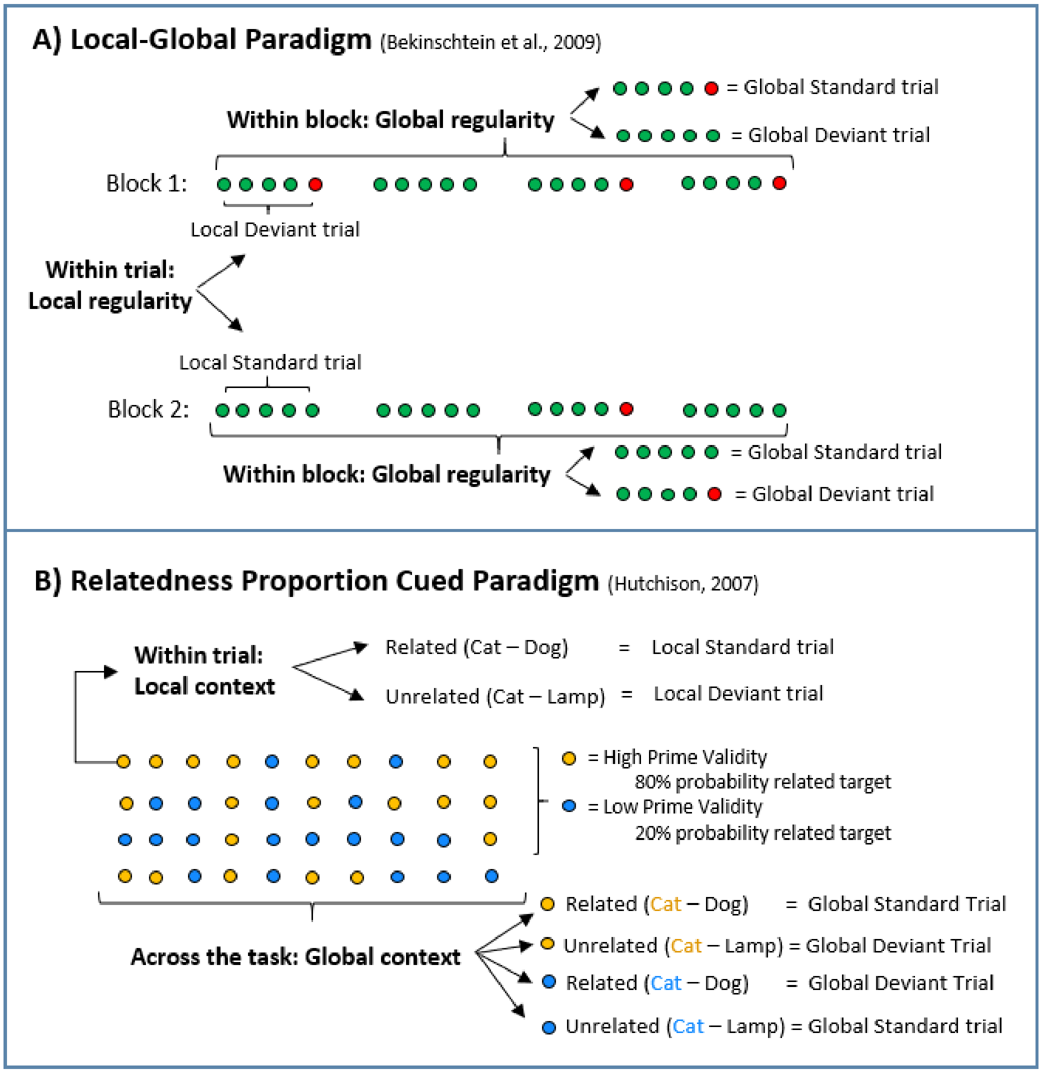


Figure 1.1: *Parallel between local-global (Bekinschtein, et al., 2009) and cued relatedness proportion (Hutchison, 2007) paradigms. In panel A, each dot represents an auditory tone, and the colour represents a different pitch. Trials can be either locally deviant or locally standard; and the context in which each trial is presented will define if they are globally standard or globally deviant. In panel B, each dot represents one trial containing a word-pair, which can be either locally standard or locally deviant; moreover, the colour of the dot represents the colour in which the prime is presented, and each colour is associated with a prime validity context. So as the local-global paradigm, depending on the context in which the prime is presented (i.e., prime validity), the target would be either globally standard or a globally deviant.*

Outline of the experimental approach

The overall aim of this thesis is to propose an electrophysiological measure that can be useful for detecting residual conscious language comprehension in patients that are diagnosed with a disorder of consciousness (DOC). The paradigm that we test intends to serve as a clinically viable tool that can be investigated in future research and eventually introduced in clinical contexts. We here present a set of experiments that follow a structured order that goes from behaviour, neural correlates, until reaching the goal of providing a proof-of-concept of an experimental manipulation that has clinical purpose for the DOC patients' population; the rationale behind each experiment will be reported throughout the thesis in each chapter.

The present chapter introduced the topic of disorders of consciousness, the current diagnostic issues and the relevant role that neuroimaging and electrophysiological methods have in this matter. Moreover, this chapter reviewed evidence in the literature about electrophysiological methods to assess the presence of conscious processing in DOC patients and addressed the importance of assessing language abilities in these patients to identify markers of consciousness. Furthermore, this chapter reviewed the role of expectations in language comprehension and how these could be experimentally manipulated to assess the presence or absence of conscious processing.

The second chapter includes two behavioural experiments where we investigate the role of conscious strategic expectations on behaviour in a relatedness proportion paradigm in a semantic priming task. The third chapter includes a visual electrophysiological experiment that intends to replicate the results of the behavioural paradigm from the second chapter and investigate the neural markers of strategic semantic expectations, by assessing the violations of semantic expectations. The fourth chapter builds up from previous results, where we adapt the experimental manipulation to meet and measure patients' abilities, and therefore propose a clinically viable tool that can be used in the context of DOC patients to measure the influence of conscious semantic expectations as an indicator of conscious processing. This auditory task is first tested on a group of healthy participants to establish the neural markers of strategic semantic expectations in this specific experimental manipulation. Moreover, we tested the task in two DOC patients (1 VS/UWS, 1 MCS), so we present their data as patient cases in order to provide a proof-of-concept for this paradigm, that we are here proposing as tool that can be used to assess strategic and conscious semantic expectations in DOC patients. The fifth chapter includes the final discussion of this set of experiments and propose future research directions for the findings here exposed.

CHAPTER 2: BEHAVIOURAL STUDIES EXPLORING SEMANTIC EXPECTATIONS

Introduction

The aim of this thesis is to investigate markers of consciously controlled semantic expectations for more accurate diagnoses in prolonged disorder of consciousness. As a first step, the present chapter intends to identify an experimental paradigm that encourages healthy individuals to use conscious and strategic semantic expectations. Thus, these measures could subsequently serve as indicators of conscious semantic processing in DOC patients.

The brain is an active processor of information and as predictive coding accounts suggest, it produces expectations at different hierarchical neural levels with varying complexity to make sense of sensory stimuli and process information about the world (Heilbron & Chait, 2018; Clark, 2013). Several authors have investigated how expectations influence the processing of language and how this human ability operates under the principles proposed by predictive coding (Lewis & Bastiaansen, 2015; Ylinen et al., 2016; Lau et al., 2013; Hutchison, 2007; Kuperberg & Jaeger, 2016). Evidence for the influence of predictions in language comprehension stems mainly from single words (Lau et al., 2013; Hutchison, 2007; Keefe & Neely, 1990) and sentence (Bonhage, Mueller, Friederici, & Fiebach, 2015; Rommers, Meyer, Praamstra, & Huettig, 2013; DeLong, Urbach & Kutas, 2005) paradigms.

The semantic priming paradigm

The semantic priming paradigm is a renowned single words experimental manipulation, in which a prime word is followed by either a related or an unrelated target word (Neely, 1991). Thus, the semantic priming effect correspond to individuals showing faster behavioural responses (i.e., RT) to target words that are preceded by a related prime in contrast to target words preceded by an unrelated prime (Neely, 1976). There are two broad categories of theories that explain how the semantic priming effect is produced (Hutchison, 2007). First, priming is considered an automatic process that occurs as a result of spreading activation (Collins & Loftus, 1975; Quillian, 1967), where concepts (representation of words) are stored in our memory system in form of interconnected nodes of information that are associated with one another (Balota & Lorch, 1986; Collins & Loftus, 1975). As language is constituted by a vast vocabulary, each concept is stored as a node within the semantic net and linked to other associated nodes (Collins & Loftus, 1975). When individuals encounter a prime word, the stored representations for that specific word are activated and this activation simultaneously propagate to semantically associated words; therefore, pre-activating the potential associated target (Klinger, Burton, & Pitts, 2000; Neely, 1991; Balota and Lorch, 1986). The view of spreading activation as an unconscious automatic process has been supported by studies that have found priming effects when presenting individuals with masked primes; which means that the prime word is presented for a brief period of time (e.g. 50ms) and is "hidden" between two mask patterns (e.g. string of letters), where individuals usually report seeing the masks but not being consciously aware of the prime's perception (Rolke et al., 2001; Kiefer & Brendel, 2006; Kiefer, 2002; Deacon et al., 2000). For example, in a masked semantic priming study, the primes were presented to individuals for only 50ms; however, the results showed priming effects in all four experiments that the authors reported, even though individuals were unaware of the prime words (Klinger et al., 2000).

Secondly, the semantic priming effect may be a result of pre-lexical expectancy mechanisms operating, as individuals engage in expectations about the target based on the prime (Hutchison, 2007; Becker, 1980). A target word can only be semantically associated to a limited amount of words; therefore, when seeing the prime, individuals may expect the target to be within those expected possible words. On the contrary, a prime can be unrelated to a vast amount of target words and thus it is not possible for the comprehender to engage in an expectation (Hutchison, 2007; Hutchison, Neely, & Johnson, 2001). Hence, individuals will show faster responses (e.g., RTs) to related targets as a result of facilitation for the target's lexical access, because the target falls within the expected potential words; while unrelated targets will show a delayed response due to inhibition, as the target fails to fulfil their expectations (Keefe & Neely, 1990). Unlike automatic priming due to spreading activation, expectancy mechanisms require consciously controlled semantic priming; however, controlled priming only occurs when the stimulus onset asynchrony (SOA) between prime and target has a sufficient length of time that allows individuals to form a conscious expectation about the target (Hutchison, 2007). Therefore, some authors have proposed that both automatic and expectancy mechanisms play a role in the semantic priming effects, and that their recruitment will rely on task demands (Neely & Keefe, 1989); for example, at shorter SOA semantic priming is thought to rely on automatic mechanisms and at longer SOA priming effects are a result of controlled expectancy mechanisms, as individuals have time to direct their attentional resources to the prime and formulate an expectation about the target (Hutchison, 2007; Hill et al., 2002).

Lexical decision task vs Pronunciation task

The lexical decision task (LDT) and pronunciation task (PT) are two ways of measuring RTs in semantic priming paradigms (Neely, 1991). In the LDT, a prime word is presented followed by either a related target word, unrelated target word or a nonword target. Individuals have to make a lexical decision as they are requested to judge whether the target is either a word or a nonword, and the RTs are extracted from these responses (Neely & Keefe, 1989). When the target is preceded by a related prime, individuals are generally faster to recognise the target as a word, than when the target is preceded by an unrelated word (Hill et al., 2002; Perea & Rosa, 2002; Neely & Keefe, 1989; Lupker, 1984). In pronunciation tasks, prime words precede either related or unrelated targets and individuals have to pronounce the target aloud as soon as they perceive it, so the onset of their pronunciation is registered as RTs (Keefe & Neely, 1990). Studies have found that individuals pronounce the target faster when it is preceded by a related prime than an unrelated prime (Hutchison, 2007; Perea & Gotor, 1997; Keefe & Neely, 1990).

In the LDT, post-lexical mechanisms are assumed to operate; once individuals encounter the target, they check backwards whether the target is related or unrelated to the prime (Keefe & Neely, 1990; Seidenberg, Waters, Sanders, & Langer, 1984). For example, an LDT study by Chwilla, Hagoort & Brown (1998) found evidence for the use of backward priming, where they used bidirectionally related (“spider”-“Web”), backward-related (e.g. baby – stork), and forward-related (e.g. mouse – cheese) word-pairs, unrelated word-pairs that served as the baseline, and nonword targets. The primes were presented auditorily and the targets visually. The behavioural results showed priming effects in all three types of related targets (bidirectional, forward, and backward) with respect to unrelated targets. The presence of semantic priming effects in backward targets indicates that individuals are checking the prime-target relatedness in a backward way, as the relationship can only be stablished from target to prime (Chwilla et al., 1998). Knowing the prime-target relatedness provides valuable information for the word/nonword response selection; if it is a related target the response can only be a word, while unrelated targets may be either a word or a nonword (Keefe & Neely, 1990). The backward priming mechanism supposes a problem as individual’s responses would be biased when making a word/nonword decision. When the target is related to the prime, individuals would be biased to respond word; on the contrary, they would be more likely to provide a nonword response when the target is not related to the prime, when in fact, the target can also be an unrelated word (Neely & Keefe, 1989; Seidenberg et al., 1984). The word or nonword bias generates either facilitation or inhibition in the behavioural responses (Neely, 1991). For example, if nonword targets are presented in a higher proportion than unrelated word targets, a nonword facilitation would be produced, as individuals would be aware that after an unrelated target it is highly likely that the response should be a nonword (Neely, 1991).

On the other hand, priming effects observed in pronunciation tasks are proposed to rely on pre-lexical mechanisms and act in a forward direction; thus, pronunciation tasks would be exempt from backward priming effects, which reinforces the use of expectancy mechanisms (Hutchison, 2007; Seidenberg et al., 1984). A study by Kahan, Neely and Forsythe (1999), gathered previous studies results where backward priming in both LDT and PT were investigated (Koriat, 1981; Peterson & Simpson, 1989; Seidenberg et al., 1984; Shelton & Martin, 1992). The authors emphasized the general finding that backward priming effects were reported in all SOA manipulations in LDTs, whilst in PTs these priming effects were only found at short SOAs (<300ms) and not at longer SOAs (>450ms). Moreover, Kahan, Neely and Forsythe (1999) measured backward priming effects at two SOAs (150ms and 500ms), by selecting word-pairs in which prime-target were not associated forwardly, but were strongly associated backwardly, either by forming a word (e.g., HOP – Bell), or semantically associated (e.g., WOOD – termite). Similar to the previous results, the authors reported backward priming effects for both SOAs in the LDT (24-30ms); whereas the results for the PT only showed significant backward priming effects for the shorter SOA (13ms), and no effect was observed in the longer SOA (4.5ms). These results provide evidence that individuals do not rely on post-lexical mechanisms in word pronunciation at longer SOAs (Kahan et al., 1999).

Priming effects reported across the literature are diverse and highly dependent on task demands, so Neely and Keefe (1989) proposed a hybrid three-process theory as they suggest that a single process theory is not able to account for the priming effects diversity. Their hybrid theory includes automatic spreading activation, expectancy, and semantic-matching processing as the mechanisms responsible to account for priming effects (Neely, 1991). The mechanisms would act according to the extent that they are required by task demands. For example, spreading activation would explain priming effects in both LDT and PT as long as the experimental manipulation includes a short SOA; expectancy mechanisms would operate in PT at long SOAs; and semantic-matching processing accounts for priming effects in the LDT when a word/nonword response is required (Neely, 1991; Neely & Keefe, 1989).

The relatedness proportion paradigm

A common behavioural experimental manipulation to test for the influence of controlled expectancy mechanisms in language processing is the relatedness proportion paradigm (RP) in a semantic priming task (Hutchison, 2007; de Groot, 1984; Neely et al., 1989; Keefe & Neely, 1990); here, experimenters manipulate the proportion of related word-pairs across experimental blocks (Neely et al., 1989; de Groot, 1984) or across the entire semantic priming task (Hutchison, 2007). Studies have shown that priming effects are greater in contexts where there is a higher proportion of related word-pairs (high validity), with respect to contexts of lower proportion of related word-pairs (low validity) (de Groot, 1984; Neely et al., 1989). The difference in the priming effects driven by low or high validity conditions is known as the relatedness proportion effect (RPE) and has been reported in studies using both LDT (Neely et al., 1989; de Groot, 1984; Perea & Rosa, 2002; Grossi, 2006) and PT (Hutchison, 2007; Keefe & Neely, 1990). In the RPE, individuals attend to the prime, from which they generate a conscious expectation about the upcoming target when in a high validity context, as they are aware that the target is likely to be related to the prime (Keefe & Neely, 1990).

For example, Hutchison (2007) used a pronunciation task to measure RPEs, where the primes were presented in one of two colours to cue participants about the prime validity conditions, with each colour representing either a high or low prime validity (high: 78%; low: 22%). Therefore, when participants encounter the prime, they would have information regarding the probability of the upcoming target being related or unrelated to the prime. The author used a short SOA (267ms) and a long SOA (1240ms) aiming to identify automatic and consciously controlled priming, respectively; and an RPE only at the long SOA. The results showed priming effects regardless of the RP manipulation, as participants responded faster to related targets relative to unrelated targets. Moreover, individuals had faster RTs for the shorter SOA; however, the priming effects were stronger at the long SOA. Regarding the RPE, the author reported a significant effect only at the long SOA. The relatedness proportion effect, especially at long SOAs, is thought to rely on consciously controlled expectancy mechanisms; therefore, individuals would require controlling their attentional resources in order to enhance task performance (Hutchison, 2007). For this purpose, the author measured participant's attentional control using a test battery composed of three tests that were completed before the relatedness proportion task: the ospan task that measures working memory by combining arithmetic and semantic information (Kane & Engle, 2003); the antisaccade task that measures cognitive inhibition where individuals are required to look away from a certain stimuli to perceive the target (Kane et al., 2001); and the Stroop task that measures selective attention and response speed by presenting written colour names with the same or different font colour (Spieler, Balota, and Faust, 1996), see Hutchison (2007) for more details. Participants were divided into groups according to their battery performance (scores), resulting in three groups: high, moderate and low attentional control. The results showed that the magnitude of the RPE grew linearly with increasing attentional control, meaning that individuals with higher scores at the attentional control battery showed stronger RPEs, relative to individuals who scored lower in the battery and showed no RPE, providing evidence that expectation generation in this paradigm is effortful (Hutchison, 2007). These results provide evidence for semantic priming as a result of both automatic (short SOA) and controlled (long SOA) semantic processes (Hutchison, 2007). Moreover, the priming effects were stronger at long SOA, meaning higher facilitation for related targets when individuals rely on controlled semantic processing. Furthermore, RPEs mean greater priming effects in contexts where there is a higher proportion of related word-pairs (high prime validity), due to facilitation produced by the contextual information. Hence, facilitation is a result of generating a conscious expectancy about the target, and this conscious expectation can only occur when individuals are given sufficient time (long SOA) to engage in a conscious expectation (Hutchison, 2007).

Nonword ratio confounder in the LDT

Studies using LDT have also reported RPEs in healthy individuals (Stolz, Besner, Carr, 2005; Neely et al., 1989). A study by Stolz, Besner, and Carr (2005) used related prime-target word-pairs and non-word target word-pairs. The authors conducted a series of experiments manipulating the RP (.25, .50 and .75) and SOA (200ms, 350ms, 800ms), which resulted in 9 groups. First of all, priming effects were reported in all nine groups, where each group had a different RP and SOA. Moreover, semantic priming effects were significantly greater in the higher RP (.50 and .75) in contrast with lower RP (.25) in the 350ms and 800ms SOA conditions, indicating the presence of RPEs. At the lowest SOA of 200ms, semantic priming effects did not significantly vary at the three different RP (Stolz et al., 2005). In an LDT, both the unrelated word targets and nonword targets have no semantic or associative relationship with the prime; thus, when individuals see an unrelated target there is a probability that the target would be a nonword, which is known as the nonword ratio (NR) (Neely et al., 1989). For example, if we have a total of 200 word-pairs, where 80 are related prime-target words, 20 are unrelated prime-target words, and 100 are prime-target nonwords, the nonword ratio would be .83.3 as 83.3% of unrelated items correspond to nonwords (Neely et al., 1989). The nonword ratio and the relatedness proportion are confounders, as the increases in the RP lead to unavoidable increases in the nonword ratio and vice versa (Keefe & Neely, 1990). By taking the nonword ratio example given above, 80 word-pairs are related and 20 word-pairs unrelated; hence, the probability of a target word being related to its prime (i.e., RP) is .80. In this particular example the NR and RP are very similar (not always the case), so by decreasing the RP to .20, the NR would decrease to .555. If we would use these values for an LDT, the RPEs would be induced by the RP manipulation, but they could also be caused by the variations in the NR. In conditions with higher NR, individuals are more likely to give nonword responses for unrelated targets, which would in turn increase the priming effects for this condition with respect to conditions with a lowest NR (Neely et al., 1989).

Pronunciation task encourages expectancy mechanisms

Pronunciation tasks (PT) have been used to assess the influence of expectancy mechanisms in RPE because individuals do not rely on semantic matching mechanisms when they are requested to pronounce the target (Hutchison, 2007). As PT does not include word/nonword responses, it produces purer RPE as these effects are not confounded with the increases in the nonword ratio (Keefe & Neely, 1990, Hutchison, 2007). Contrary to LDT, in a PT knowing if the target is related or unrelated to the prime does not provide relevant information for a pronunciation response (Keefe & Neely, 1990). For example, Keefe & Neely (1990) used two relatedness proportions to compare two groups of participants that performed a pronunciation task, where one group had high prime validity (RP .88) and the other group low prime validity (RP .33). Moreover, the related word pairs belonged to the same category (e.g. prime: BIRD), although some targets corresponded to a high-dominance (e.g. robin) or low dominance (e.g. swan). In a previous study Neely, Keefe and Ross (1989) used the high and low dominance word-pairs because from an expectancy mechanisms perspective, high-dominance targets have a higher probability to be part of the expectancy set of words that individuals generate when perceiving the prime, than the low-dominance targets that are less likely to be considered as an expected target (Neely et al., 1989). Therefore, in the pronunciation study (Keefe & Neely, 1990) the authors anticipated that if individuals are using expectancy mechanisms, the results would show significant RPE only for high-dominance targets and no significant RPE for low-dominance targets. The results yielded a significant RPE (between high .88 and low .33 prime validity groups) of 18-19ms for high-dominance targets and no significant RPE (1-2ms) for low-dominance targets;

hence, providing evidence for the generation of expectations from the prime in this pronunciation manipulation (Neely et al., 1989). Hence, RPE in pronunciation tasks are proposed to be produced by expectancy mechanisms, where individuals first generate an expectancy about the target based on the prime, and at the same time generate a conscious strategic expectancy that relies on the context (high-low validity) in which the word-pair is presented (Hutchison, 2007; Keefe & Neely, 1990). The expectancy mechanisms act to facilitate semantic processing of words; thus, successful expectations provide faster behavioural responses at two levels: at a local level, faster responses to related than unrelated targets (SPE) and at a global level, higher priming effects in context with high validity relative to low validity (RPE) (Hutchison, 2007; Keefe & Neely, 1990).

Outline of the experimental approach

In order to investigate the influence of strategic conscious expectations in language processing, we perform a partial replication of a relatedness proportion manipulation in a semantic priming task implemented by Hutchison (2007). As mention above, the author used a pronunciation task and presented the word-pairs using both a short (267ms) and long SOA (1240). Hutchison (2007) made a review of previous RP studies that reported RPEs (see Hutchison (2007) table 1), which used either an LDT or PT, and SOAs ranging from 45ms to 1200ms. The author found that the RPE magnitude increases when increasing SOA; studies reported negative RPE under 100ms, and only positive RPE from 200ms; however, only after 400ms the RPEs grow considerably, and this SOA has been previously attributed as a sufficient time to allow for the emergence of strategic priming (Neely, 1977; Hutchison, 2007). Thus, for a pronunciation task to recruit strategic priming, the SOA has to be long enough to allow participants to engage in conscious expectations about the target based on both the prime and the prime validity (Hutchison, 2007). For the previous reasons, an SOA of 400ms was the minimum length to consider for the current experiment; in addition, RPEs elicited by longer SOAs in RP manipulations using pronunciation tasks, specifically 1000ms (Keefe & Neely, 1990) and 1240ms (Hutchison, 2007), were already reported in the literature. Thus, we consider an SOA of 800ms as it was only previously reported in a study using an LDT task (Stolz et al., 2005), but not previously reported in a pronunciation task. Moreover, we selected this specific manipulation for the following reasons that were reviewed above: pronunciation tasks rule out the NR and RP confounder, as participants are not requested to perform a word/nonword response; pronunciation tasks are known to rely on pre-lexical mechanisms that invoke expectancy mechanisms, contrary to LDT where individuals rely on post-lexical semantic-matching mechanisms. Moreover, for the purpose of this thesis, it is not within our interests to measure attentional control with a test battery as it was implemented by Hutchison (2007); therefore, we will only conduct the cued relatedness proportion priming task reported by the author.

The aim of the current study is to generate a local and a global semantic context that allows participants to generate local and global controlled conscious expectations as it was implemented in the local-global paradigm (Bekinschtein et al., 2009). For this purpose, we provide a local semantic context at shorter time-scales (within trial) using word-pairs, where individuals can generate local expectations about the target based on the prime. In addition, we simultaneously create a global semantic context at a longer time-scale (across the task) given by the RP manipulation, in which individuals can use the prime cue (i.e., colours) to engage in a global expectation about the likelihood of the target being related to the prime. We expect to find semantic priming effects as a result of facilitation given by local expectations, where we will observe faster responses (RT) to related targets relative to unrelated targets. Moreover, we hypothesised that participants would show an interaction between the validity of the prime and the relatedness of the target, meaning greater priming effects under a high prime validity context, in contrast with a low validity context. The interaction would indicate the use of the global context to facilitate processing of words, as individuals would show more facilitation in a highly valid context.

Behavioural Experiment 1

Methods

Participants

Participants were recruited through the Research Participation Scheme website of the University of Birmingham and received credits for their participation. We recruited 32 participants, as it was indicated by the number of possible counterbalancing permutations; however, the data of one participant was excluded from the analysis as their reaction times were classified as an outlier by the non-recursive procedure for outlier elimination (detailed below; Van Selst & Jolicoeur, 1994; Hutchison, 2007). Hence, the sample size for the present experiment includes 31 participants (27 female and 4 male; median age: 19; age range: 18-26). All participants reported to be mono-lingual native English speakers, meaning that they learned English from birth within the United Kingdom; right-handed; and had no history of neurological conditions or diagnosis of dyslexia. All participants gave written informed consent prior to their participation in this study that was approved by the Science, Technology, Engineering and Mathematics Ethical Review Committee of the University of Birmingham, UK (ERN_15-1367P).

Stimuli

The present study is a partial replication of the relatedness proportion manipulation from Hutchison (2007). We selected associated prime-target word pairs from the Semantic Priming Project database (Hutchison et al., 2013), which is an available database composed of 1661 prime-target word pairs with their respective item characteristics, such as Forward associative strength (FAS), which is to the proportion of individuals who spontaneously named the same target word after reading the prime word; moreover, the letter length of both prime and target (length); the frequency (Subfreq) of the word within the vocabulary of both prime and target; the orthographic neighbourhood (OrthoN) that refers to how many words can be formed if we change one letter as the other letters maintain their order (Hutchison et al., 2013).

We ordered the word-pairs in the database by FAS and selected the first 360 word-pairs with higher values, so the stimuli contain word-pairs where the prime and target have a high forward association. As the database was created in the United States, we eliminated 8 word-pairs due to cultural differences as the current study was conducted in the United Kingdom (e.g., Clorox-bleach; slacks-pants); therefore, the stimuli for this experiment was composed of 352 associated prime-target word-pairs, where the 156 first word-pairs that had the higher FAS values were labelled as critical stimuli and were used for the statistical analysis. Furthermore, the remaining 196 word-pairs served as fillers to create the prime-validity manipulation across the task, and these were not included in the statistical analysis. We manually divided the 156 critical word-pairs into two lists (N:78 word-pairs per list) and moved word-pairs until the lists were balanced by forward association, length, log HAL frequency, and orthographic neighbourhood according to the values specified in the database (all $p > 0.604$; all $BF_{10} < 0.196$; see table 2.1). Next, we divided all 196 filler word-pairs into two balanced lists using the same values as above (N: 98 word-pairs per list; all $p > 0.284$. all $BF_{10} < 0.267$, see table 2.1).

Critical Related (1,2)	**t**	**df**	**P**	**BF_{10}**
FAS	-0.31	154	0.757	0.18
Prime Length	-0.402	154	0.689	0.186
Prime Frequency	0.029	148	0.977	0.176
Prime OrthoN	-0.062	148	0.951	0.176
Target Length	0.443	154	0.658	0.189
Target Frequency	-0.47	154	0.639	0.191
Target OrthoN	-0.519	154	0.605	0.195
Filler Related (1,2)	**t**	**df**	**P**	**BF_{10}**
FAS	0.109	194	0.913	0.156
Prime Length	0.735	194	0.463	0.2
Prime Frequency	-0.03	193	0.978	0.156
Prime OrthoN	-0.75	193	0.452	0.203
Target Length	0.753	194	0.452	0.202
Target Frequency	1.07	194	0.285	0.266
Target OrthoN	-0.686	194	0.494	0.193

Table 2.1: *Independent samples T-tests and Bayesian Independent Samples T-test to test for significant differences between lists of word-pairs (Critical related 1,2; Filler related 1,2) using categories specified in the semantic priming project database (Hutchison et al., 2013).*

To create the unrelated word-pairs lists, we manually repaired all word-pairs in each of the four lists mentioned above (two critical, two filler); specifically, we shuffled the targets words, while the primes remained in the same position. Next, we checked that unrelated targets were semantically unrelated to their prime word. Overall, we created eight lists: two critical related, two critical unrelated, two filler related, and two filler unrelated. Each participant was assigned two critical lists (one related and one unrelated; 78 word-pairs per list), and two filler list (one related and one unrelated; 98 word-pairs per list). Thus, each participant saw all words within the full set of 352 word-pairs exactly once, where half were related word-pairs and the other half unrelated word-pairs.

To create the prime-validity manipulation we first assigned half of the critical word-pairs, including both related and unrelated items, to one colour (yellow or blue), and the other half with the other colour in an interleaved order. Next, the related filler set was assigned with one colour (yellow or blue), and the unrelated filler set was assigned with the other colour. Therefore, across all items seen by each participant, 77.8% of word-pairs presented in one of the two colours were related, thus giving that colour a high prime validity, and 77.8% of word pairs presented in the other colour were unrelated, hence giving that colour low prime validity. Across participants, the colour assignment of the high validity primes was counterbalanced (i.e., half of participants saw high prime validity word-pairs in blue and low prime validity word-pairs in yellow; and the other half saw the opposite colours for each proportion), and all possible combinations of word lists were used, resulting in 32 permutations.

Design and procedure

We presented the task with Psychtoolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) in Matlab (Mathworks, 2016). Participants performed a pronunciation task, and their responses were recorded with a microphone, where each recording lasted 2.5 seconds. Each trial started with a black fixation cross in a grey background lasting 600ms; followed by the prime word displayed at the centre of the screen in either yellow or blue for 160ms; next, a blank screen remained for 640ms, and then the target was presented at the centre of the screen in black (stimulus onset asynchrony, SOA: 800ms). The target word stayed on the screen for 2500ms while the microphone recorded the pronunciation responses. Next, a blank screen was presented for 1000ms as the interstimulus interval (ISI) preparing the participant for the next trial. The trial procedure is shown in figure 2.1. Participants had breaks every 10 minutes.

Each participant was tested individually and sat approximately 70 cm away from a laptop screen. All participants received written information about the study, the experiment instructions, and the consent form. The instructions were repeated by the experimenter before starting the experiment, followed by four practice trials. Participants were instructed that a coloured uppercase word (either blue or yellow) will be displayed on the screen and that they must read it silently to themselves; then, a black lowercase word will be displayed on the screen, and they should pronounce the word aloud, as fast and accurately as possible. Participants were told that the colour of the uppercase word will cue the probability of the lowercase target being related or unrelated. Half of the participants received the following written instructions: “If the uppercase word is Blue, it is highly likely that the meaning of the lowercase word will be related; and if the uppercase word is Yellow, it is highly likely that the meaning of the lowercase word will be unrelated" (as per Hutchison, 2007). The other half of participants received the same instructions but with the opposite colours.

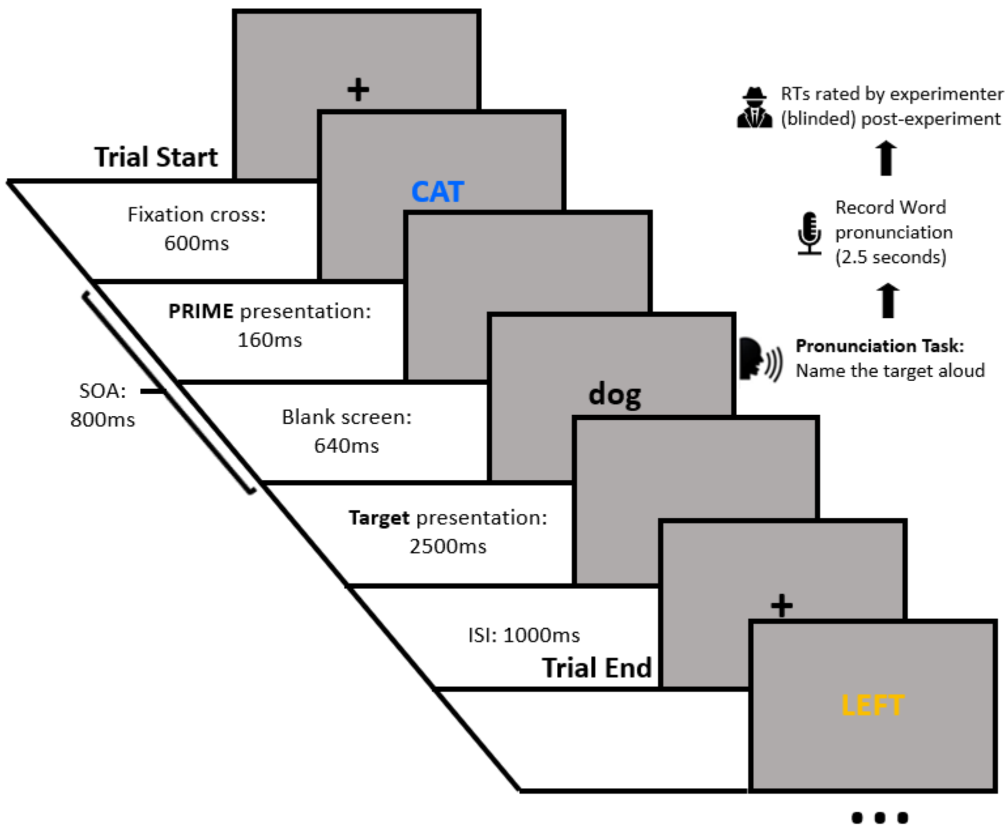


Figure 2.1: *Experimental design Behavioural Experiment 1. Semantic Priming Relatedness Proportion task (Hutchison, 2007). Participants were required to name the target word aloud and as fast as possible, while their responses were recorded with a microphone.*

Data analyses

Once the full sample of participant's data was collected, the experimenter manually identified the word onset in milliseconds using Audacity Software (Audacity Team, 2014) to obtain reaction times (RTs), where a total of 4,992 audio files were analysed (critical items from 32 participants). The condition in which each word was presented was blinded to the experimenter to avoid subjective bias in RTs.

For the statistical analyses of RTs, we discarded trials where the recording was unintelligible, and trials in which participants did not provide a response. Measuring RTs aims to test the speed with which participants respond to the different experimental conditions; however, raw reaction times tend to be skewed as a result of outliers, which could influence the mean and provide an inaccurate estimate of the participants' performance (Lo & Andrews, 2015). One solution for this problem is to make the data normal by log transforming it, but this loses crucial information about the response speed (Lo & Andrews, 2015), directly affecting the main aim of the task. Therefore, we chose to eliminate outlier RTs following the same procedure as in Hutchison (2007), which is the non-recursive procedure for outlier elimination (Van Selst & Jolicoeur, 1994). Specifically, the number of trials from each condition (of each participant) represented the sample size and the average was calculated using the RTs from each condition (condition mean). Reaction times that were more than X standard deviations from the condition mean were considered to be outliers and were removed, where the value of X decreases with decreasing sample size and is anchored at X=2.5 for a sample size of 100. Furthermore, we applied the procedure across participants to determine whether any participant was an outlier with respect to others, and one participant was classified as an outlier. Hence, we analysed the data of 31 participants, where a median of 38 trials (range: 35-39) comprised the high related condition; a median of 38 trials (range: 35-39) the high unrelated condition; a median of 38 trials (range: 35-39) the low related condition; and a median of 38 (range: 36-39) the low unrelated condition.

We conducted all behavioural analyses in JASP 0.8.3.1 software (JASP Team, 2017). First, we tested for an interaction effect between the relatedness of the target and the validity of the prime on reaction times. Thus, we conducted a two-way repeated measures ANOVA with factors of relatedness (i.e., related vs unrelated targets) and prime validity (i.e., high vs low prime validity). Moreover, we reported equivalent Bayesian Repeated Measures ANOVAs (Van Doorn et al. 2019; Wagenmakers et al., 2018). We hypothesized that individuals would show faster RTs for related (expected) word-pairs with respect to unrelated (unexpected) targets as a result of facilitation given by semantic association (i.e., semantic priming effect). In addition, we hypothesized individuals would provide evidence of an interaction consisting of larger difference between related and unrelated word-pairs in a high validity context, in contrast to a low validity context. In other words, we expected larger priming effects in the high validity context than in the low validity context, as a consequence of participants using the validity of the prime to expect the likelihood of the target being related to the prime.

Results

The two-way repeated measures ANOVA analysis yielded a significant interaction between the relatedness of the target and the validity of the prime ($F(1, 30) = 4.437$, $p=0.044$, $\eta^2= 0.129$), shown in figure 2.2. Moreover, the Bayesian Repeated Measures ANOVA ($BF_{inclusion} = 1.765$) showed weak/anecdotal evidence in support of the alternative hypothesis, which states that there is an interaction between prime validity and target relatedness. Furthermore, the interaction is a result of significantly different reaction times to related items in high and low contexts ($F(1,30) = 6.368$, $p = 0.017$), whereas reaction times to unrelated targets in both high and low contexts were similar ($F(1, 30) = 1.778$, $p = 0.192$). Moreover, there was a significant difference between related and unrelated targets in both high prime validity context ($F(1, 30) = 70.22$, $p < 0.001$) and low validity context ($F(1, 30) = 99.12$, $p < 0.001$), see table 2.2.

Condition	Low Validity = 22.2% **Mean RTs (SD)**	High Validity = 77.8% **Mean RTs (SD)**	**Prime Validity Effect**
Unrelated	465ms (56ms)	459ms (50ms)	
Related	437ms (55ms)	422ms (53ms)	
Priming Effect	28ms (25ms)	37ms (16ms)	9ms (24ms)

Table 2.2*: Descriptive statistics including Mean RT (ms) and standard deviation of related and unrelated word-pairs on each validity context, High Prime Validity and Low Prime Validity. Semantic priming effects and prime validity effect (relatedness proportion effect).*

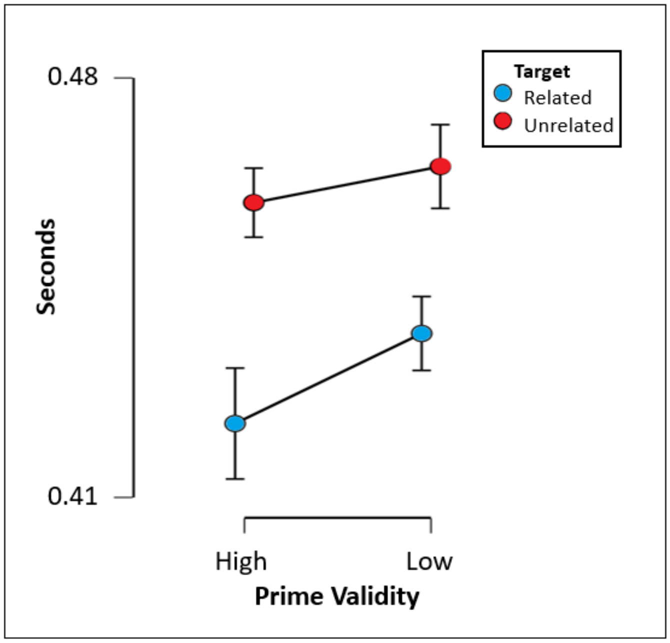


Figure 2.2*: Mean RTs, CI: Prime Validity (High / Low), Relatedness of the target (Related / Unrelated).*

Interaction (p = 0.044) between the validity of the prime and the relatedness of the target.

Discussion Behavioural Experiment 1

Expectancy mechanisms are considered to be involved in language processing, and a way of investigating this influence is using a relatedness proportion paradigm (RP) in a semantic priming task. In the present study, we conducted a partial replication of the Cued-RP Priming task that was implemented by Hutchison (2007), where we also used a pronunciation task; however, our SOA of 800ms was slightly shorter than the SOA of 1240ms used by the author. As expected, our results shows significant priming effects (28ms low prime validity; 37ms high prime validity), where individuals pronounce the related targets faster than unrelated targets. Furthermore, we observe a significant interaction between the validity of the prime and the relatedness of the target (9ms RPE), which reflects greater priming effects under a high validity context in contrast to a low validity context.

Regarding Hutchison's (2007) study, we replicated their results as we both found semantic priming effects and relatedness proportion effects in a pronunciation task at long SOA. The author reported an RPE at the longer SOA (1240) of 9+-8ms across all participants, and we are reporting an RPE of 9ms using an SOA of 800ms; moreover, our Bayesian analysis showed anecdotal evidence in support of the interaction. However, the author found an interaction of RP with attentional control (AC) level, meaning that the magnitude of the RPE was modulated by the level of AC, which was obtained through the attentional control battery tasks. Individuals that fell within the high AC level, showed an RPE of 19ms; while moderate AC level presented an RPE of 8ms; and an RPE of 1ms for low AC individuals

Hutchison (2007). In our study, attentional control was not measured as in Hutchison (2007); hence, we cannot differentiate between high, moderate and low AC. By not measuring this variable (AC) we do not know which percentage of our participants would have been considered as a high AC or low AC, and our effect –similar to moderate AC- could be the result of averaging participants from all three categories.

Although we replicated Hutchison's (2007) Cued-RP Priming task by reporting RPEs in a group of 31 participants, our observed interaction was not statistically strong ($p=0.044$) according to our alpha level ($p<0.05$) for hypothesis testing. A weak behavioural effect supposes a risk in accomplishing the general aim of the present thesis, which is testing a paradigm that measures strategic semantic expectations, so these can be then assessed in patients diagnosed with disorders of consciousness using electrophysiological measures. For this paradigm to become clinically viable, we would expect a strong behavioural effect, so its electrophysiological markers could be then investigated in patients. Therefore, having a strong and reliable behavioural effect becomes an essential preliminary step before proceeding to electrophysiological measures. Here, our results yield a behavioural effect that has a strong support in the literature, although is not strong enough to become clinically feasible; thus, we next conducted a second behavioural study to investigate whether we can obtain a stronger effect by changing some experimental parameters. The factors that could have caused the weak RPE and the proper experimental adjustments for the next experiment are discussed below.

The first question that arises from a weak RPE is whether selecting different stimuli would increase the strength of our RPE. Several studies have reported that normatively associated prime-target word-pairs (e.g., SPIDER-web) elicit stronger priming effects, in contrast with word-pairs that only hold a semantic relationship (e.g., SPIDER-ant) (Cruse et al., 2014; Plaut, 1995, Shelton & Martin, 1992). Therefore, in Experiment 1, we selected prime-target word pairs from a validated associative word-pairs databased by Hutchison and colleagues (2013). Our main selection criteria was the Forward association strength (FAS) indicator, although our stimuli lists were also balanced for other indicators provided in the database (i.e., length, frequency, orthoN). As previously mentioned, the FAS corresponds to the proportion of individuals that mention a certain target when they hear a specific prime; for example, if 80% of participants mention 'DOG' (target) when they hear the word 'CAT' (prime), the FAS for that word-pair is 0.8 (Hutchison et al, 2013; Hutchison, 2003). All the associated word-pairs that we selected for the present experiment (352), correspond to those with the highest FAS in the entire database, including the filler word-pairs. However, the critical items -those included in the statistical analyses- have higher FAS (Mean = 0.69, SD = 0.09) than the filler word-pairs (Mean = 0.47, SD = 0.05); and the lists of word-pairs were balanced to ensure participants were assigned with lists composed of similar features (see table 2.1 methods section). The present RPE is weak but the priming effects are statistically strong ($p<0.001$) as participants were significantly faster to name related targets relative to unrelated targets. The presence of priming effects is an indicator that the selected stimuli has the required features to elicit priming effects and eventually RPE. Therefore, our weak RPE may be caused by other factors than our stimuli selection from the normative associated database (Hutchison et al., 2013).

Another question is whether we could detect a stronger effect if we also included measures of attentional control (AC) as in Hutchison (2007), who used a battery of tasks to measure attentional control, where individuals were classified as high, moderate, or low AC according to their test scores. Our RPE was similar to the RPE for the moderate AC category in the author's results, possibly our results fall in the middle AC category, as our general RPE of 9ms represents the average of participant's RPEs. Our effect could have followed the same trend if we would have measured AC in the same way as the author, as only 58% (18/31) of our participants showed positive RPEs, while the other 42% presented negative RPEs. Therefore, we include in the next behavioural experiment a self-report form, where we ask participants after they complete the task, to report whether they were following the task and whether they were using a strategy while performing it (for more details see methods section of behavioural experiment 2 methods section). However, it is not in our interest to measure attentional control per se as in Hutchison (2007), as we aim at to exploring participant's performance in the RP task itself, and not measure AC using independent tests. In our current manipulation we instructed participants to pronounce the target aloud and to follow the rule, which stated the colour that was highly likely to be related or unrelated, but we did not measure whether individuals were actually following the task, being expectant to the target when hearing the prime, or using the colour cue of the prime; therefore, the self-report form focuses on these aspects about the task.

A different question to explain the weak effect is whether a different SOA manipulation could increase the strength of our effect. We test an SOA of 800ms that has not been previously reported in the literature in a RP manipulation using a pronunciation task, only in an LDT (Stolz et al., 2005), as we intended to provide new evidence for that specific SOA. There is no clarity whether this SOA could have influenced the detection of a weak RPE; nevertheless, as our aim is the detection of a strong effect reflecting the use of conscious semantic expectations, in the next experiment we will test the same SOA of 1240ms as Hutchison (2007). The increase in SOA is based on the view that providing participants with longer time between prime and target reinforces the use of strategic processing (Hutchison, 2007; Keefe & Neely, 1990).

To further ensure obtaining a stronger RPE in the next experiment and ensure that the effect exists, we will increase the sample size. Button and colleagues (2013) proposed that when studies reach weak effects but within the alpha level ($p<0.05$) for hypothesis testing, which is the case of the current study ($p=0.04$), and the study is replicated using the same sample size, the replication study will only have 50% power at $p<.05$ to detect the same sized effect (Button et al., 2013). Consequently, we are doubling the original sample of 31 participants, as a way to increase the power to estimate the true effect size. Thus, the sample size for behavioural experiment 2 is increased to 62 participants.

As the overall aim of the present thesis is to find a marker of conscious strategic expectations in language processing, it becomes absolutely necessary to have a strong behavioural effect that serves as a basis to explore subsequent electrophysiological markers; as tools to assess awareness in DOC patients using electrophysiological techniques should be based and supported by underlying behavioural effects, detected through objective measures, strategic control and/or subjective measures such as self-report (Seth et al., 2008). Therefore, next we report a second behavioural study, in which we replicate the current study, with the changes described above. Moreover, as patients' EEG assessments should be carried out as single subjects, it is crucial to provide evidence of a robust effect on a group level.

Behavioural Experiment 2

This study and the visual EEG study in Chapter 3 are published in one article in the journal Eneuro, therefore some parts match with those of the publication: Vidal-Gran, C., Sokoliuk, R., Bowman, H., and Cruse, D. (2020) Strategic and non-strategic semantic expectations hierarchically modulate neural processing, *Eneuro, 7(5).*

Methods

Participants

We recruited participants through the Research Participation Scheme website of the University of Birmingham, who received credits for their participation. As the previous behavioural study had a sample size of 31 participants, we decided to double the sample size in order to have more power to detect the effect; thus, we recruited a total of 64 participants. We then excluded the data of two participants from analysis due to outlying data, as quantified by the non-recursive procedure for outlier elimination (Van Selst & Jolicoeur, 1994; Hutchison, 2007), same procedure used in Experiment 1. Therefore, the final sample consisted of 62 participants (59 female, 3 male; median age: 19, range: 18 – 28). All participants reported to be mono-lingual native English speakers, right-handed, and with no history of neurological conditions or diagnosis of dyslexia. All participants gave written informed consent prior to participation in this study, which was approved by the STEM Ethical Review Committee of the University of Birmingham, England.

Stimuli

The stimuli used in the present experiment are the same as Experiment 1.

Procedure

The task was presented with Psychtoolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) in Matlab (Mathworks, 2016). The procedure is similar to behavioural Experiment 1; although there are some differences as they were explained above (discussion behavioural experiment 1). We increase the SOA between prime and target from 800ms to 1240ms in this experiment and vocal reaction times (RTs) are measure with a Cedrus SV-1 Voice Key (Cedrus Corporation). These measures are less time consuming with respect to manual rating of word recordings and are commonly used in pronunciation studies (Hutchison, 2007; Moritz & Graf, 2006; Moritz et al., 2001; De Houwer, Hermans, & Spruyt, 2001; Pecher, Zeelenberg, & Raaijmakers, 1998; Balota & Chumbley, 1985). The use of voice key to obtain pronunciation onset RTs could introduce voice key bias, which means that the word onset response times can vary due to each word having different initial letters that generate different sounds (Kessler, Treiman, & Mullennix, 2002). However, in our study (within-subjects) all participants are exposed to the same words and have to pronounce all the words from the stimuli set, so these differences should average out when comparing RTs across participants. Furthermore, we include a self-report form at the end of the experiment to analyse whether participants were consciously and strategically following the task (more details below). All participants completed four practice trials under the experimenter´s supervision to adjust the voice key threshold according to the participant´s speech volume.

Each trial starts with a central fixation cross on a grey background lasting 600 ms; then, the prime word was displayed in either yellow or blue, at the centre of the screen for 160 ms; followed by a blank screen for 1080ms, and subsequently the target was displayed on the screen; thus, the stimulus onset asynchrony (SOA) was 1240ms. The target stayed on the screen until the participant pronounced the word; then the word disappeared from the screen, which remained blank for 300ms. Afterwards, a rating for the quality of pronunciation was displayed on the screen with the following questions and potential responses: How would you rate your pronunciation? 1) Correct pronunciation; 2) Unsure of pronunciation; 3) Mispronunciation; 4) Accidental voice-key triggering. Participants gave a button response on the keyboard (1-4) to rate their pronunciation (as per Hutchison, 2007). After the participant responded, the screen remained blank for 1000ms (ISI), before the next trial began, the trial procedure is shown in Figure 2.3.

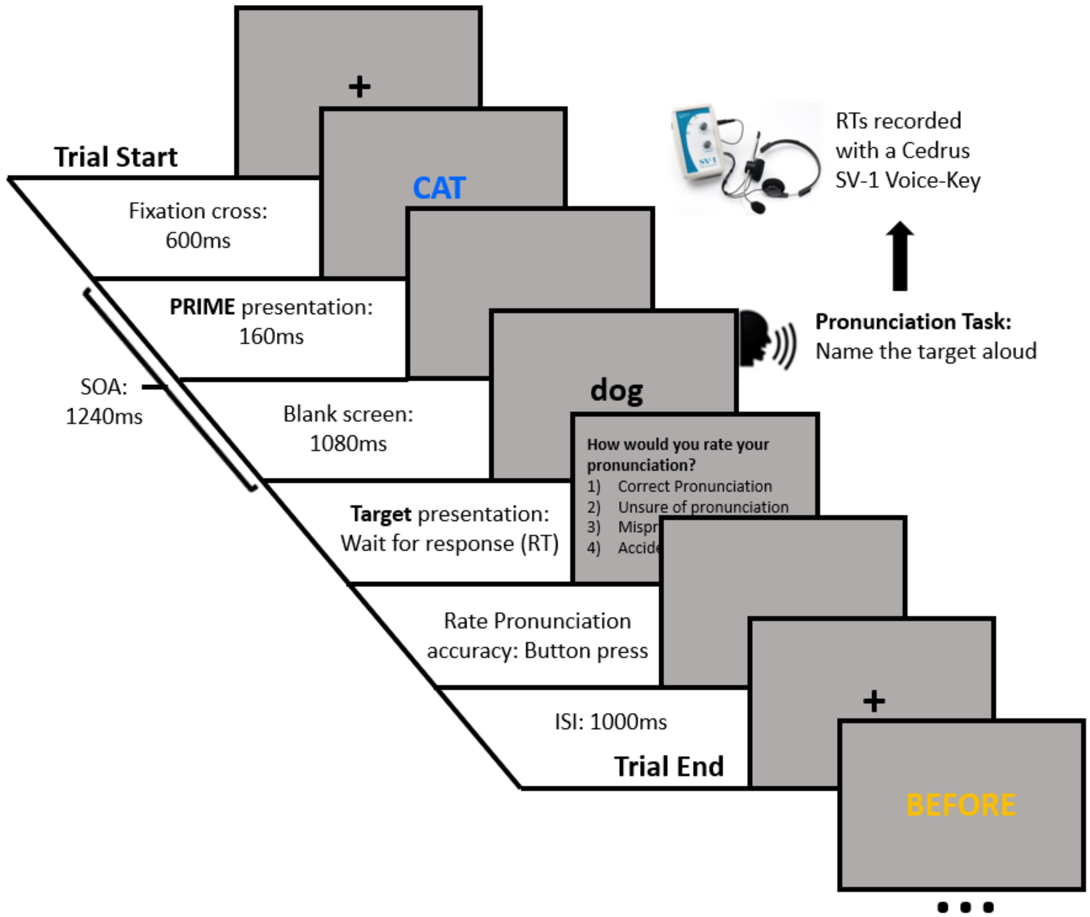


Figure 2.3: *Experimental design Behavioural Experiment 2. Semantic Priming Relatedness Proportion task (Hutchison, 2007). Participants were required to name the target word aloud and as fast as possible, while their responses were recorded.*

Participants completed the task in a sound-attenuated cabin, which isolated external noise that could have triggered the voice-key microphone. The instructions provided were also the same as experiment 1, where we instructed participants that a coloured uppercase word (either blue or yellow) will be displayed on the screen and that they must read it silently to themselves; then, a black lowercase word will be displayed on the screen, and they should pronounce the word aloud, as fast and accurately as possible. Participants were told that the colour of the uppercase word will cue the probability of the lowercase target being related or unrelated. Half of the participants received the following written instructions: “If the uppercase word is Blue, it is highly likely that the meaning of the lowercase word will be related; and if the uppercase word is Yellow, it is highly likely that the meaning of the lowercase word will be unrelated” (as per Hutchison, 2007). The other half of participants received the same instructions but with the colours flipped.

After the task, we asked participants to complete a self-report form about the use of strategy throughout the task, to determine whether they were using expectations strategically. The form was composed of three questions and a free text description of the strategy. The questions were the following: 1) Which colour was highly likely to be related? (Responses: BLUE / YELLOW); 2) Did you use the colour of the UPPERCASE word (BLUE, YELLOW) as a cue for knowing whether the following word was related or unrelated? (Responses: YES / NO); 3) Did you engage in any strategy to speed up your responses using the colour cue? (Responses: YES / NO); 4) If YES, briefly describe. We considered participants to have used strategic expectation (i.e., those referred to subsequently as the Strategy group) if they correctly identified the colour that was assigned for the high validity condition

(Question 1), answered YES in questions 2 and 3, and described a strategy in question 4. All other participants were classified into the No Strategy group.

Behavioural Data Analyses

To ensure the inclusion of trials pronounced correctly, we only included trials that were rated by the participants with a correct pronunciation (button press 1); moreover, we eliminated RTs that were longer than 2500ms and shorter than 1ms (i.e. not correctly triggered by the vocal onset). To eliminate outlier trials, we also chose (as in Experiment 1) to follow the non-recursive procedure for outlier elimination (Van Selst & Jolicoeur, 1994) as in Hutchison (2007). Next, across all participants we used the same procedure to determine outlier participants and rejected data from two participants that met the outlier criteria. For the remaining 62 participants, a median of 37 trials (range: 16-39) contributed to the high related condition; a median of 36 trials (range: 12-39) to the high unrelated condition; a median of 37 trials (range: 16-39) to the low related condition; and a median of 36 (range: 15-39) contributed to the low unrelated condition.

All behavioural analyses were conducted in JASP 0.8.3.1 software (JASP Team, 2017). To test for an effect of global context on reaction times, we conducted a two-way repeated measures ANOVA with factors of relatedness (i.e., related vs unrelated targets) and prime validity (i.e., high vs low prime validity). We also reported equivalent Bayesian Repeated Measures ANOVAs (Van Doorn et al., 2019; Wagenmakers et al., 2018). We expected individuals to show faster RTs for related (expected) in contrast with unrelated (unexpected) targets due to local level expectations – i.e., priming. Furthermore, we expected an interaction, with larger priming effects in a high validity context in contrast with a low validity context, reflecting the use of global level context to predict upcoming stimuli. As a follow-up analysis, we conducted a three-way ANOVA, with its Bayesian equivalent, to test for the interaction and the report of strategy vs no strategy (self-report form) as a between-subjects factor.

Results

In a two-way repeated measures ANOVA, we found a significant interaction between prime validity and relatedness of the target ($F(1, 61) = 13.751$, $p < 0.001$, $\eta^2 = 0.184$), which was also supported by a Bayesian Repeated Measures ANOVA ($BF_{inclusion} = 19.25$). As shown in Table 1, this interaction stems from the larger semantic priming effect in the high prime validity context ($F(1, 61) = 42.58$, $p < 0.001$) relative to the low prime validity context ($F(1, 61) = 26.72$, $p < 0.001$). Furthermore, reaction times to unrelated items were markedly similar across contexts ($F(1, 61) < .001$, $p = 0.999$), while the difference in semantic priming stems from significantly different reaction times to related items ($F(61) = 14.421$, $p < 0.001$), see Table 2.3.

Condition	Low Validity = 22.2% **Mean RTs (SD)**	High Validity = 77.8% **Mean RTs (SD)**	**Prime Validity Effect**
Unrelated	508ms (76ms)	508ms (75ms)	
Related	493ms (73ms)	472ms (76ms)	
Priming Effect	15ms (32ms)	36ms (54ms)	21ms (60ms)

Table 2.3*: Descriptive statistics including Mean RT (ms) and standard deviation of related and unrelated word-pairs on each validity context, High Prime Validity and Low Prime Validity. Semantic priming effects and prime validity effect (relatedness proportion effect).*

Of 62 participants, 32 were classified in the "No-Strategy" group and 30 were classified in the "Strategy" group. A post-hoc mixed design ANOVA with two within factors (Relatedness of Target; Validity of the prime) and one between subjects factor (Strategy; No-strategy) revealed a significant Target * Prime Validity * Strategy interaction ($F (1, 60) = 7.537$, $p=0.008$, $\eta^2= 0.090$, $BF_{inclusion} = 3.203$), reflecting the apparent presence of a prime validity effect when participants reported using the prime strategically ($F (1, 29) = 20.388$, $p < 0.001$, $\eta^2= 0.413$; $BF_{inclusion} = 34.67$) but absence of a prime validity effect when participants reported no strategy ($F (1, 31) = 0.860$, $p = 0.361$, $\eta^2= 0.027$; $BF_{inclusion} = 0.393$; Figure 2.4). The No strategy group, however, did exhibit a significant semantic priming effect by showing faster responses in the related relative to unrelated items ($F (1, 31) = 21.656$, $p < 0.001$, $\eta^2= 0.411$; inclusion $BF_{inclusion} = 4994.57$).

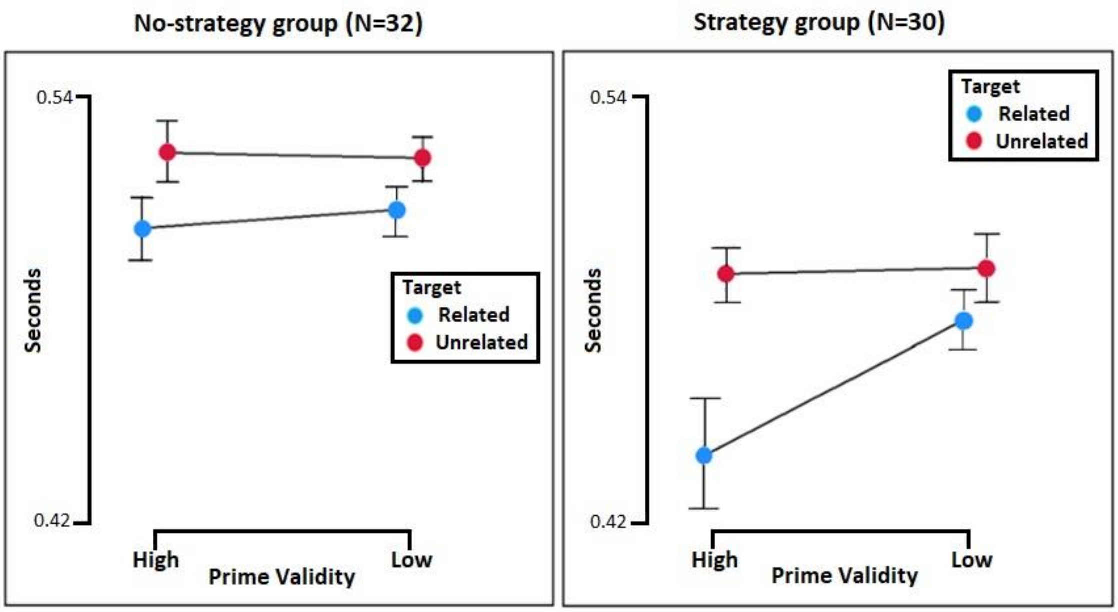


Figure 2.4*: Mean RTs / CI: Prime Validity (High / Low), Relatedness of the target (Related / Unrelated). Interaction (p < 0.001) between the validity of the prime and the relatedness of the target in the group of participants that reported the use of a conscious strategy (right), and no interaction (p = 0.361) in the group of participants that did not report a conscious strategy (left).*

Discussion Behavioural Experiment 2

The overall aim of this thesis is to provide evidence for the use of conscious controlled semantic expectations in language comprehension and their respective electrophysiological markers, which can be used to assess these abilities in noncommunicative patients. Therefore, the aim of this chapter is to find a strong behavioural effect that could account for the use of strategic expectations when processing language stimuli. For this purpose, we use the local-global paradigm rationale by measuring violation of local expectations on a within-trial level accounting for non-strategic processing; and violations of global expectations across the experiment that evidence the use of strategic processing (Bekinschtein et al., 2009). The relatedness proportion paradigm fits within this rationale, as on a local level, unrelated targets constitute a violation of the local expectations that are only produced for related targets, and on a global level the use of the prime validity (e.g., colour cue) across the task allows the generation of strategic expectations involving the global context in which stimuli is presented (Hutchison, 2007; Keefe & Neely, 1990). Here, a violation of global expectation is an unrelated target from a high validity prime, or a related target from a low validity prime

In the present experiment we are reporting significant priming effects that indicate that on a local level, individuals are responding faster to expected targets (related) relative to unexpected targets (unrelated). Moreover, on a global level we report significant relatedness proportion effects, which indicate larger priming effects in a high validity context relative to a low validity context, as individuals engage in more complex global predictions that result in facilitation for related words presented under a highly predictable context (Boudewyn, Long, & Swaab, 2015). The current results replicate Behavioural Experiment 1, however, showing a stronger RPE, which was the aim for repeating this experimental manipulation. The stronger effect, with respect to behavioural experiment 1, could be caused by the extension of the SOA between prime and target from 800ms to 1240ms, as individuals have more time to generate a controlled expectation about the target (Hutchison, 2007), and splitting participants by the use of strategy with the self-report form, provides evidence that participants require to actively follow the task in order to show RPE; moreover, the increased sample size provides greater confidence that the observed effect is close to the true effect size in the population (Button et al., 2013).

In Behavioural Experiment 2 we include a self-report form that is completed after performing the task, and individuals were classified according to their responses (see method section for more details) either into the strategy group or non-strategy group. Out of 62 participants, both groups are reasonably even as 32 participants are classified into the non-strategy group and 30 participants into the strategy group. The follow-up analysis yielded that individuals that reported the explicit use of strategy while performing the task show significant RPE, whereas for individuals that report no strategy involvement there are no significant RPEs. However, the non-strategy group showed significant priming effects, indicating that they were only engaging in local semantic expectations and not applying conscious global expectations. Strategic control can be understood as individuals using their resources and the information provided by a certain task to have a successful performance (Seth et al., 2008); specifically, in this task individuals use the colour cue to know the probability of a target being related and unrelated. A complex level of processing, in which individuals are capable to use the task demands to their advantage, indicates both the presence of conscious processing and the use of a strategy to maintain task demands and remember instructions (Seth et al., 2008).

Together, these behavioural data are consistent with a dissociation between a local expectation about the identity of the target generated by the prime, and a global expectation about the relatedness of the target that necessitates reportable, effortful, and strategic application of expectation (Vidal-Gran, Sokoliuk, Bowman and Cruse, 2020; Hutchison, 2007; Lau et al., 2013). In the current experiment we provide evidence for the use of strategic semantic expectations in this paradigm, which can be justified by the following reasons. First, relatedness proportion manipulations encourage the use of strategic priming as individuals use the context in which stimuli are presented, which can be blocks or across the task (Hutchison, 2007; Keefe & Neely, 1990). Moreover, using a pronunciation task also encourages the generation of expectations about the target from the prime, ruling out the nonword ratio confounder, as individuals are not requested to perform a decision response on whether the target is a word or a nonword as in studies using an LDT (Neely et al., 1989). Hence, individuals only rely on forward word-pair association and not semantic matching mechanisms, as knowing the semantic relationship between prime and target does not provide any relevant information for the pronunciation response (Keefe & Neely, 1990). Next, presenting prime and target word-pairs with a long SOA (1240ms) provides sufficient time for individuals to generate conscious strategic expectations about the target involving the global context, as expectations that require the use of contextual information are more complex and more time consuming than automatic or less strategic processes operating (Hutchison, 2007). In addition, the finding indicating that only individuals that report the use of a conscious strategy show a global effect reflects direct evidence for the use of strategic processing while performing the task. For all the reasons stated above, the current paradigm seems successful in the detection of strategic semantic expectations; therefore, in the next chapter we report the same experimental manipulation while we measure individuals' electrophysiological recordings, to investigate the neural markers of strategic semantic expectations in this task. Although recording EEG activity when participants are moving (i.e., pronouncing words) supposes a significant challenge because of muscle artefact created by the movement (Fargier, Bürki, Pinet, Alario, & Laganaro, 2018), the pronunciation task is included in the next experiment to allow direct comparison between behavioural measures and electrophysiological markers of strategic semantic expectations, which are directed towards a future clinical application.

CHAPTER 3: ELECTROPHYSIOLOGICAL MARKERS OF STRATEGIC SEMANTIC EXPECTATIONS IN A VISUAL RELATEDNESS PROPORTION PARADIGM

Behavioural Experiment 2 and aspects of this visual EEG study are published in one article in the journal Eneuro, therefore some parts match with those of the publication: Vidal-Gran, C., Sokoliuk, R., Bowman, H., and Cruse, D. (2020) Strategic and non-strategic semantic expectations hierarchically modulate neural processing. *Eneuro, 7(5).*

Introduction

In the previous chapter, we explore a behavioural paradigm in healthy participants to investigate the influence of strategic semantic expectations in language comprehension, as the overall aim of this thesis is to propose an electrophysiological tool that could serve to detect residual language comprehension in patients diagnosed with disorders of consciousness. The main feature of this potential tool should be the distinction between non-strategic and conscious strategic processing, as it is achieved in the local-global paradigm, which has been broadly tested in DOC patients research (Sergent et al., 2017; Faugeras et al., 2012; Bekinschtein et al., 2009; El Karoui et al. 2015). In addition, the tool should include language stimuli that can account for the presence of language processing, as understanding language is a crucial ability in determining an individual's level of awareness (Owen & Coleman, 2008). Therefore, in the previous Chapter we propose that the relatedness proportion manipulation by Hutchison (2007) follows the same rationale as the local-global paradigm, by presenting word-pairs (related and unrelated) that can generate expectations in shorter time-scales within-trial; and simultaneously presenting the word-pairs within a global context, which is the prime validity manipulation (coloured primes) that cue participants about the likelihood of the target being related, which encourages the use of global expectations that are built at longer time-scales across the task (Hutchison, 2007; Bekinschtein et al., 2009). Our previous results provide behavioural evidence for the use of global strategic semantic expectations in this paradigm, that are generated in a top-down manner as they involve the use of contextual information (e.g., observed in strategy group), as opposed to low-level expectations that can be generated from the local context when no active global strategy is taking place (e.g., observed in no strategy group).

In the present experiment we investigate the neural correlates for the use of strategic semantic expectations in this relatedness proportion paradigm. According to our previous results, only participants that report using a strategy while performing the task show a behavioural relatedness proportion effect (RPE), which is the effect that we expect to find on a neural level to evidence semantic strategic expectations. Hence, in this study we only include individuals that are assigned to the Strategy group as a result of their strategy self-report. In the local-global paradigm, the presence of a global effect (i.e., P3B component) in DOC patients reflects conscious processing (Bekinschtein et al., 2009); therefore, in this experiment we are interested in observing how the use of global expectations behaves on a neural level by maintaining the same experimental parameters as our previous studies, so that in an upcoming experiment we can adapt the task to meet patients' abilities.

As I previously reviewed, language processing at a neural level is thought to be influenced by the recruitment of expectancy mechanisms which are supported in predictive coding accounts (Lewis & Bastiaansen, 2015; Ylinen et al., 2016; Kuperberg & Jaeger, 2016). Predictive coding theory argues that the brain processes information in a hierarchical probabilistic Bayesian manner (Friston, 2005; Knill & Pouget, 2004) by contrasting sensory input with prior expectations generated from context and the perceiver's knowledge (Heilbron & Chait, 2018; Clark, 2013). Expectations are sent down from higher levels of the hierarchy and any subsequent unexplained sensory input is sent back up the hierarchy as prediction error (Heilbron & Chait, 2018; Friston & Kiebel, 2009; Rao & Ballard, 1999).

Some argue that evoked neural responses (e.g., event-related potentials [ERPs]) reflect prediction errors (Chennu et al., 2013; Friston, 2005). For example, the Mismatch Negativity (MMN) is larger in amplitude for stimuli that do not match short-term auditory expectations, relative to those that do (Heilbron & Chait, 2018). Prediction errors at higher levels of the hierarchy are investigated in paradigms that introduce violations of expectations formed from the global context in which stimuli occur. Indeed, generating such expectations involves complex cognition including working memory and report of conscious expectation (e.g., Bekinschtein et al., 2009). The local-global paradigm (Bekinschtein et al., 2009) elegantly pits local expectation within each trial (i.e., standard vs deviant pitch tones) against a global expectation built from the context across blocks of trials. This paradigm elicits an initial MMN to local violations of expectation, and a subsequent centro-parietal positivity at approximately 300ms post-stimulus (P3b) to global violations of expectation (see Faugeras et al., 2012; King et al., 2014; El Karoui et al., 2015); thereby, separating prediction error signals at two levels of an expectation hierarchy that unfold sequentially.

Within the realm of more ecologically valid stimulus processing, speech comprehension is similarly influenced by expectations at multiple levels of a hierarchy (e.g., Lewis & Bastiaansen, 2015; Ylinen et al., 2016; Lau et al., 2013; Hutchison, 2007; Kuperberg & Jaeger, 2016). The N400 – a negative deflection peaking around 400ms post-stimulus (Kutas & Federmeier, 2011) – is a potential marker of errors of such semantic expectations (Rabovsky & McRae, 2014). On a local level, the N400 is larger to words that have not been primed relative to those that have (e.g., larger for DOG when preceded by Lamp than by Cat; Cruse et al., 2014; Lau et al., 2013; Koivisto & Revonsuo, 2001), and at a more global level, the N400 is larger to words that are unexpected within a sentential context (Brothers, Swaab, & Traxler, 2017; Boudewyn et al., 2015; Thornhill & Van Petten, 2012; Van Berkum, Hagoort, & Brown, 1999). Interestingly, unlike the MMN/P3b in auditory processing, semantic prediction errors appear to be reflected in the magnitude of a single component –the N400– rather than in a series of components moving through the hierarchy of relative top-down involvement.

One approach to separate prediction error signals at two levels of a semantic expectation hierarchy is with a prime validity manipulation of a word-pair priming task. Specifically, we can pit the facilitation of target word processing that comes from presentation of a related prime against a global context in which it is not efficient for the comprehender to use the prime to predict the target – i.e., primes rarely followed by related targets (Keefe & Neely, 1990; Hutchison, 2007; Lau et al., 2013(a); Lau et al., 2013(b)). Therefore, as the proportion of related pairs increases within a context, the prime validity increases (i.e., the prime is more likely to predict the target). If individuals use the global context of prime validity to modulate their expectations, behavioural facilitation follows.

In ERP studies of prime validity, this hierarchy of local expectations (i.e., the prime relatedness) and global expectations (i.e., the prime validity) has not been reported to modulate the amplitudes of two sequential components (Boudewyn et al., 2015; Lau et al., 2013); hence, there is no evidence of a two-stage profile to semantic expectation violation. Rather than reflecting error at one level, the N400 (or see Boudewyn et al. (2015) for N200 evidence) appears to account for a combination of errors across levels of the hierarchy. To disentangle these results, here we report a pre-registered trial-by-trial manipulation of both local and global semantic expectations. First, we report a replication of the reaction time facilitation caused by global context as described by Hutchison (2007). Second, we report the associated electrophysiological markers of expectation and violation across levels of the hierarchy from a separate group of healthy participants performing the same task. In accordance with predictive coding, we hypothesised that ERP amplitudes would reflect violations of expectation at consecutive levels of the hierarchy, with local violations evident earlier than global violations.

EEG Visual Experiment

Methods

This study was pre-registered in the Open Science Framework website and details can be found under the following link: https://osf.io/npvby. Any deviations from the pre-registered methods and analyses are specifically stated in the text.

Participants

We recruited participants through the Research Participation Scheme website and placed advertisement posters at the University of Birmingham; participants received a monetary compensation for their participation. We recruited 37 participants, however, since we only investigated those who reported using a strategy, the final sample only included 22 participants (15 female, 7 male; median age: 21, range: 18 - 30; classified by the same report form as experiment 1). The inclusion criteria were the same as those for Experiment 1; however, participants were also required to attend for a structural T1-weighted MRI scan at the University of Birmingham Imaging Centre (BUIC); therefore, participants who had any metal parts in their body, were claustrophobic, or women who were pregnant were excluded from the study, as the scan was mandatory for participation. All participants gave written informed consent prior to participation in this study, which was approved by the STEM Ethical Review Committee of the University of Birmingham, England.

We aimed to detect a reaction time interaction of the same magnitude as seen in the Strategy group of Experiment 1; therefore, we conducted a power analysis to select an appropriate sample size for this goal. We performed non-parametric power calculations using the data of all participants of the Strategy group from Experiment 1. Specifically, from the pool of participants of the Strategy group, we selected with replacement N participants and conducted the same two-way repeated measures ANOVA 1000 times to test for the reaction time interaction effect. With an N of 22 participants in the Strategy group we achieved 80% power at $p<.05$ (i.e., 80% of ANOVAs included a significant interaction).

As we did not know if a participant was in the Strategy group until their self-report form was completed at the end of the study, we recruited participants until 22 of them were classified as being in the Strategy group (median age: 21, range: 18-30; 12 in the no-strategy group, median age: 22, range: 19-33). After removal of trials rated as mispronunciations and those considered outliers according to the non-recursive outlier elimination procedure of Van Selst & Jolicoeur (1994; as Experiment 1 and 2), a median of 28 trials (range: 11-38) contributed to the high related condition; a median of 29.5 trials (range: 13-38) to the high unrelated condition; a median of 29 trials (range: 12-39) to the low related condition; and a median of 28 (range: 14-37) contributed to the low unrelated condition.

Stimuli and procedure

The stimuli were the same as experiment 1 and 2; however, we corrected for unrelated targets that had overlapping phonemes with their respective related target to avoid word confounders. We placed primes, related, and unrelated targets in three columns, followed by both targets translated into phonemes and we manually checked that there were no overlapping phonemes between related and unrelated targets; when we found overlapping phonemes, we swapped unrelated targets within-list, so the lists remained balanced as specified in Table 1.

The procedure was the same as in Experiment 2, except for the duration of the fixation cross that is increased from 600ms to 750ms to provide more time for an EEG time-frequency baseline; and the target remaining in the screen for 2500ms. See trial procedure in Figure 2.3.

EEG recording

The EEG signal was continuously recorded with a 125 channel AntNeuro EEG system (AntNeuro b.v., Enschede, Netherlands) at a sampling rate of 500 Hz, with impedances kept below 20 kΩ. We placed the ground electrode on the left mastoid bone and referenced online to CPz. As participants were required to pronounce words aloud, we also recorded a bipolar EMG signal with one EMG electrode above the upper lip and the other below the lower lip on the left side of the mouth; approximately over the superior and inferior Orbicularis Oris muscles (Lapatki, Stegeman & Jonas, 2003; Drake, Vogl & Mitchell, 2009).

EMG Pre-processing

As this task involved participants speaking, there were considerable artefacts in the EEG data around the vocal reaction time that were challenging to remove adequately. We therefore chose to analyse only the EEG data up to the point of vocal artefact. To minimise artefacts from additional preparatory muscular activity prior to vocal onset, in our pre-registered methods, we planned to choose the latest time-point for analysis post-target by identifying when the mouth EMG signal began to significantly differ between prime validity conditions in a temporal cluster mass randomisation test, as implemented in Fieldtrip (Oostenveld et al., 2011). However, this approach revealed no significant clusters (smallest cluster $p = 0.513$), and so did not provide a suitable cut-off time-point for our analyses. Therefore, in a deviation from the pre-registered plan, we chose our latest time-point of EEG data to analyse as 150ms prior to the fastest mean RT across conditions (in this instance High Validity – Related = 532ms; see Kuperberg, Delaney-Busch, Fanucci, & Blackford, 2018, for a similar approach). Our post-target time-window therefore continued to 382ms post-target. From all the trials included for the statistical analysis only 5.76% of trials had RTs earlier than this time-point, comparable with previous studies (Kuperberg et al., 2018).

EEG Pre-Processing Pipeline

We low pass filtered the continuous EEG data at 40Hz using the finite impulse response filter implemented in EEGLAB (Delorme & Makeig, 2004). Due to our interest in analysing slow-waves (see below), we performed no high-pass filtering. Next, we segmented the filtered EEG signals into epochs from 750ms before the onset of the prime up to 382ms post-target (see above for details). Subsequent artefact rejection proceeded in the following steps based on a combination of methods described by Nolan, Whelan, & Reilly (2010) and Mognon, Jovicich, Bruzzone, & Buiatti (2011).

First, as in the behavioural data analysis, we excluded all trials in which the participant rated their response as incorrect (i.e., 2, 3, 4 button press) and those that had reaction times that were classified as outliers in the Non Recursive Procedure for outlier elimination (Selst & Jolicoeur, 1994). Next, bad channels were identified and removed from the data. We considered a channel to be bad if its absolute z-score across channels exceeded 3 on any of the following metrics: 1) variance of the EEG signal across all time-points, 2) mean of the correlations between the channel in question and all other channels, and 3) the Hurst exponent of the EEG signal (estimated with the discrete second order derivative from the Matlab function wfbmesti). After removal of bad channels, we identified and removed trials containing non-stationary artefacts. Specifically, we considered a trial to be bad if its absolute z-score across trials exceeded 3 on any of the following metrics: 1) the mean across channels of the voltage range within the trial, 2) the mean across channels of the variance of the voltages within the trial, and 3) the mean across channels of the difference between the mean voltage at that channel in the trial in question and the mean voltage at that channel across all trials. After removal of these individual trials, we conducted an additional check for bad channels, and removed them, by interrogating the average of the channels across all trials (i.e., the ERP, averaged across all conditions). Specifically, we considered a channel to be bad in this step if its absolute z-score across channels exceeds 3 on any of the following metrics: 1) the variance of voltages across time within the ERP, 2) the median gradient of the signal across time within the ERP, and 3) the range of voltages across time within the ERP.

To remove stationary artefacts, such as blinks and eye-movements, the pruned EEG data was subjected to an independent component analysis with the runica function of EEGLAB. The Matlab toolbox ADJUST (Mognon et al., 2011) subsequently identified which components reflect artefacts on the basis of their similarity to stereotypical spatio-temporal patterns associated with blinks, eye-movements, and data discontinuities, and the contribution of these artefact components was then subtracted from the data. Next, we interpolated the data of any previously removed channels via the spherical interpolation method of EEGLAB and re-referenced the data to the average of the whole head.

Before proceeding to group-level analyses, single-subject averages for the ERP analysis were finalised in the following way. First, a robust average was generated for each condition separately, using the default parameters of SPM12. Robust averaging iteratively down-weights outlier values by time-point to improve estimation of the mean across trials. As recommended by SPM12, the resulting ERP was low-pass filtered below 20Hz using a FIR filter (again, with EEGLAB's pop_neweegfilt), and the mean of the baseline window (-200 – 0 ms) was subtracted.

Single-subject data for the time-frequency analysis were pre-processed in a similar way. However, first, we concatenated the individual trials into a matrix of channels x all time-points and filtered each channel in two-steps (high-pass then low-pass) to retain the frequency bands of interest (i.e., 8-12Hz alpha, and 13-30Hz beta), using EEGLAB's finite impulse response filter (function: pop_eegnewfilt). Next, we extracted the squared envelope of the signal (i.e., the squared complex magnitude of the Hilbert-transformed signal) to provide a time-varying estimate of power within that frequency band. The resulting time-course was re-segmented into its original epochs and averaged within each condition separately using SPM12's robust averaging procedure. As with the ERP analyses, we low-pass filtered the resulting average time-series below 20Hz (EEGLAB's pop_neweegfilt). Finally, we converted the power estimates to decibels relative to the mean of the baseline window (-200 – 0 ms.).

EEG / MRI co-registration

We recorded the electrode locations of each participant relative to the surface of the head using a Xensor Electrode Digitizer device and the Visor2 software (AntNeuro b.v., Enschede, Netherlands). Furthermore, on a separate day, we acquired a T1-weighted anatomical scan of the head (nose included) of each participant with a 1mm resolution using a 3T Philips Achieva MRI scanner (32 channel head coil). This T1-weighted anatomical scan was then co-registered with the digitised electrode locations using Fieldtrip.

Analyses

Behavioural Data Analysis:

The behavioural analyses were the same as for the Strategy Group in Behavioural Experiment 2.

EEG Analysis:

Target ERP, Prime ERP and Prime time frequency analyses:

Time-courses (ERPs / time-frequency) within the time-window of interest (0-1240ms for primes; 0-382ms for targets) were compared with the cluster mass method of the open-source Matlab toolbox FieldTrip (Oostenveld et al., 2011). This procedure involves an initial parametric step followed by a non-parametric control of multiple comparisons (Maris & Oostenveld, 2007). Specifically, we conducted two-tailed dependent samples t-tests at each spatio-temporal data-point within our time-window of interest. Spatiotemporally adjacent electrodes (t-values) with p-values < 0.05 were then clustered based on their proximity, with the requirement that a cluster must span more than one time-point and at least 4 neighbouring electrodes, with an electrode's neighbourhood containing all electrodes within an approximately 4-cm radius (median: 8, range:2-10). Finally, we summed the t-values at each spatio-temporal point within each cluster. Next, we estimated the probability under the null hypothesis of observing cluster sum Ts more extreme than those in the experimental data - i.e., the p-value of each cluster. Specifically, Fieldtrip randomly shuffles the trial labels between conditions, performs the above spatio-temporal clustering procedure, and retains the largest cluster sum T. Consequently, the p-value of each cluster observed in the data is the proportion of the largest clusters observed across 1000 such randomisations that contain larger cluster sum Ts. As our analyses were two-tailed, we set the family-wise error corrected cluster alpha to .025.

Prime slow wave linear fit analyses:

To further test for ERP evidence of expectation formation in response to the prime, we analysed whether a slow wave differentiates high validity and low validity conditions. For this comparison we used a least-squares linear fit to the averaged ERPs of each condition (High and Low validity primes) for each electrode and participant (as per Chennu et al., 2013). Next, the slope values were compared between conditions with the spatial cluster mass analysis in FieldTrip (Oostenveld et al., 2011).

Source estimation analysis:

We constructed individual boundary element head models (BEM; four layers) from subject-specific T1-weighted anatomical scans, by using the 'dipoli' method of the Matlab toolbox FieldTrip (Oostenveld et al., 2011). Next, we aligned the electrode locations, that were recorded with Xensor Electrode Digitizer device, to the surface of the scalp layer that was segmented from the T1-weighted anatomical scan. For reference points, we used the fiducial points and electrode locations as head shape. We visually checked that the electrode positions and the scalp surface were aligned, and we manually fixed imperfections. We prepared the EEG data before subjecting it to statistical analyses, where we balanced the number of trials in each condition, by taking the smallest condition N as a reference and randomly discarding trials from the other conditions surpassing that N, resulting in equal datasets.

ERPs whole brain

For the whole brain ERP source analysis, we used single-trial data that had not been subjected to robust averaging, and defined trials as time windows from -382 to 382ms relative to target onset. This data was then band-pass filtered between 1 and 40Hz using a firws filter as implemented in Fieldtrip (Oostenveld et al., 2011). Subsequently, relative to the different conditions, data were divided into seven sets: one containing all trials, one containing only related trials, one only unrelated trials, one all high-validity related and one all low-validity related trials, one containing all high-validity unrelated and one all low-validity unrelated trials. The sensor covariance matrix was estimated for all these sets of data in the time window -382 – 382ms relative to target onset. A common spatial filter was then computed on the dataset containing all trials using a Linear Constraint Minimum Variance (LCMV) beamformer (Van Drongelen, Yuchtman, Van Veen, & Van Huffelen, 1996; Van Veen, Van Drongelen, Yuchtman, & Suzuki, 1997; Robinson, 1999). Beamformer parameters were chosen including a fixed dipole orientation, a weighted normalisation (to reduce the center of head bias), as well as a regularisation parameter of 5% to increase the signal to noise ratio (cf. Popov, Oostenveld, & Schoffelen, 2018; Sokoliuk et al., 2019). This common spatial filter served then for source estimation of the remaining six sets of trials. Subsequently, the dipole moments of the different source estimates were extracted within the post-stimulus time windows of interest (time windows for source estimates of related vs. unrelated trials: 226-280ms; 232-290ms; 306-382ms; 316-350ms; time window to test interaction effect for source estimates of highly related and unrelated trials and low related and unrelated trials: 316-350ms) and their absolute values averaged over time to obtain one average source estimation value per grid point (VE) and condition.

To test for significant differences between conditions we conducted five contrasts as mentioned above; first, an interaction between prime validity (High/Low) and relatedness of the target (Related/Unrelated) in a time-window from 316 to 350ms; next, we tested the early and late main effects of relatedness of the target (Related/Unrelated) as observed in the sensor analyses results (four main effects), in their respective time windows for the early effect (226-280ms and 232-290ms); and the late effect (306-382ms and 316-350ms). Montecarlo Cluster-based permutation tests were computed as implemented in Fieldtrip (Oostenveld et al., 2011) by using averaged data over each time-window; moreover, we used an alpha and a cluster alpha level of 0.025 and 1000 permutations.

Automated Anatomical Labelling (AAL) analysis:

We tested for the post-target interaction, between the relatedness of the target (related/unrelated) and the validity of the prime (High prime validity/Low prime validity) in five specific anatomical regions of interest that are defined using the automated anatomical labelling (AAL) atlas (see Brookes et al., 2016; Sokoliuk, Calzolari, & Cruse, 2019, for similar analyses with MEG and EEG data). The selected regions are the Left inferior frontal gyrus (LIFG), including pars opercularis, pars triangularis and pars orbitralis; the posterior Left middle temporal gyrus (LMTG); and posterior Left superior temporal gyrus (LSTG), as Weber, Lau, Stillerman, & Kuperberg (2016) reported a relatedness proportion interaction in these regions. In addition, we tested the post-target interaction in the anterior LMTG and anterior LSTG, as Lau et al. (2014) found differences in the anterior left superior temporal region (LSTG) in related vs. unrelated items in a high validity condition. Moreover, as a deviation from our preregistered analyses, we tested the main effects found in the Related – Unrelated contrast at the sensor level (ERPs) in the same anatomical regions (more details in results section). To determine both the anterior and posterior parts of the LMTG and LSTG, we calculated the centre of mass of each AAL region and selected all virtual electrodes that were anterior or posterior to the centre of mass.

We aggregated the AAL regions of interest to each participant's T1-weighted image; next, for each participant individually, we extracted the average source estimation values of all VEs (from prior source estimation (cf. *ERPs whole brain*)) within each AAL region, weighted them according to their Euclidian distance to the centre of mass of the AAL region (Brookes et al., 2016) and averaged over VEs within each AAL region of interest. We then conducted paired-sample t-tests between the post-target conditions (SP-High validity / SP-Low validity) for all AAL regions; and another paired-sample t-test between the relatedness conditions (Related / Unrelated) for each AAL region in four time windows (226-280ms; 232-290ms; 316-350ms; 306-382ms) from the main effects obtained in the sensor level ERP analyses (results section). The p-values that we obtained were corrected for multiple comparisons across AAL regions using False Discovery Rate, FDR (Yekutieli & Benjamini, 1999). Furthermore, to test for evidence for the null hypothesis, we calculated Bayes

Factors using the Bayes equivalent t-test, according to Rouder, Speckman, Sun, Morey, and Iverson (2009).

Results

Behavioural Results

These results were qualitatively consistent with those we observed in Experiment 2. A two-way repeated measures ANOVA analysis showed a significant interaction between prime validity and relatedness of the target ($F(1, 21) = 9.071$, $p = 0.007$, $\eta^2 = 0.302$), while the Bayesian Repeated Measures ANOVA analysis showed anecdotal evidence for the interaction ($BF_{inclusion} = 2.519$). The interaction was driven by a larger semantic priming effect in the high prime validity context ($F (1) = 18.094$, $p < 0.001$) than in the low prime validity context ($F (1) = 4.184$, $p = 0.054$), see table 3.1. There was no significant difference between the reaction times to unrelated items across contexts ($F (1) = 0.535$, $p = 0.473$) as opposed to a significant difference between related items across contexts ($F (1) = 7.394$, $p = 0.013$).

Condition	Low Validity = 22.2% **Mean RTs (SD)**	High Validity = 77.8% **Mean RTs (SD)**	Prime Validity Effect
Unrelated	576ms (92ms)	582ms (87ms)	
Related	560ms (107ms)	532ms (110ms)	
Priming Effect	16ms (54ms)	50ms (69ms)	34ms (95ms)

Table 3.1*: Descriptive statistics including Mean RT (ms) and standard deviation of related and unrelated word-pairs on each validity context, High Prime Validity and Low Prime Validity. Semantic priming effects and prime validity effect (relatedness proportion effect).*

EEG Results – Sensor Level

Prime analyses: ERPs, time frequency and slow wave linear fit analyses

As the global context was instantiated by the prime words, we sought to also investigate potential electrophysiological markers of expectation setting (rather than post-target prediction errors). However, none of our pre-registered analyses in the prime time-window (0-1240ms after prime onset) revealed evidence of markers of expectation in response to the prime. Specifically, there were no effects in analysis of the ERPs (smallest cluster $p = 0.233$, see Figure 3.1), the slow wave linear fit analysis (no clusters formed), or the alpha-beta time-frequency analysis (smallest cluster $p = 0.136$).

Therefore, in exploratory analyses, we focused the time-window of interest for the ERP analysis on the peak of the global field power (530-1240ms), however this also revealed no significant difference between the high and low validity contexts (smallest cluster $p = 0.139$). Similarly, we used the window of interest for the alpha-beta time-frequency analysis to the peak of the global field power (602-1240ms), which also yielded no significant difference between conditions (no clusters formed). Moreover, as alpha-beta frequency bands include a wide range of frequencies we analysed them separately. However, the time-frequency analysis in the Alpha band (8-12Hz) showed no significant differences between conditions in the 0-1240ms time window (smallest cluster p = 0.121), nor in the 530-1240ms time window (smallest cluster p = 0.08). The same was true for the Beta band (13-30Hz; 0-1240ms cluster p = 0.312; 530-1240ms cluster p = 0.197). Together, these analyses suggested no apparent electrophysiological markers of pre-target expectation formation in our data.

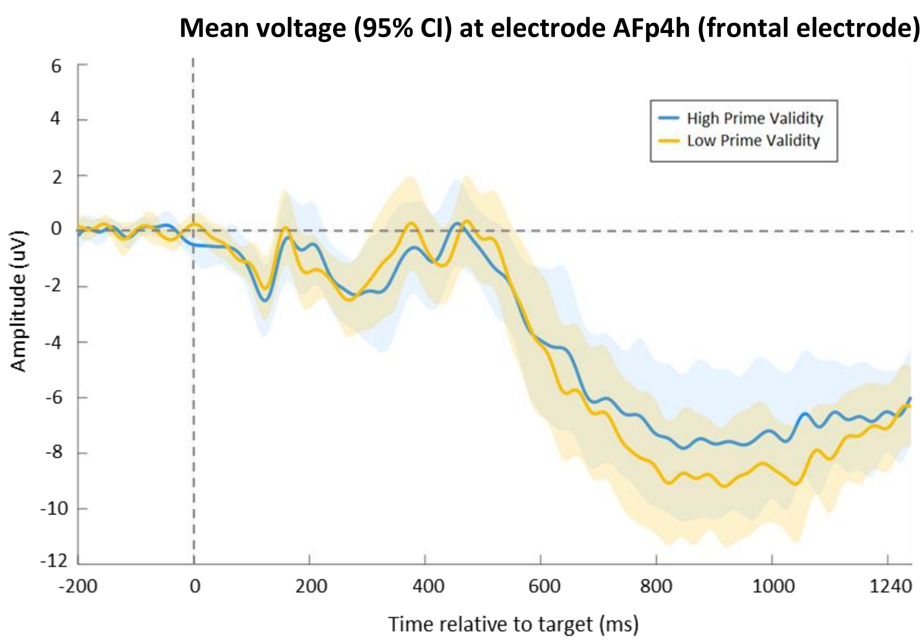


Figure 3.1: *ERP prime analysis (high vs low prime validity). ERPs containing Mean voltage (95% CI) at electrode AFp4h (frontal electrode), contrast between high and low validity primes at a time window from -200 to 1240ms post prime and pre-target, revealed no significant difference between conditions (smallest cluster p = 0.233).*

Target Results: ERPs

In our pre-registered interaction contrast in the latency range from 0 to 382ms post-stimulus, the cluster-based permutation analysis yielded no clusters. However, in pre-registered analyses of main effects in the same latency range, we found four significant main effects of relatedness of the target (i.e., unrelated versus related targets; see Figure 3.2). The clusters in our data occurred in two distinct periods within the time window as shown in Figure 3.2. Specifically, two clusters reflected a left fronto-temporal dipolar effect of relatedness (Panels A & B in Figure 3.2) at approximately 250ms post-stimulus (negative cluster: 226 – 280ms, $p = 0.019$; positive cluster: 232 – 290ms, $p = 0.009$), and two clusters reflected a later parieto-occipital dipolar effect of relatedness (Panels C & D in Figure 3.2) at approximately 350ms post-stimulus (negative cluster: 316 – 350ms, $p = 0.021$; positive cluster: 306 – 382ms, $p = 0.004$). The early effects showed a predictive signal as in both clusters the voltage exhibited more extreme values for unrelated than related items. On the contrary, the later effects showed signs of an apredictive signal, especially in Panel D, as the voltage within the cluster had more extreme values for the related relative to the unrelated items.

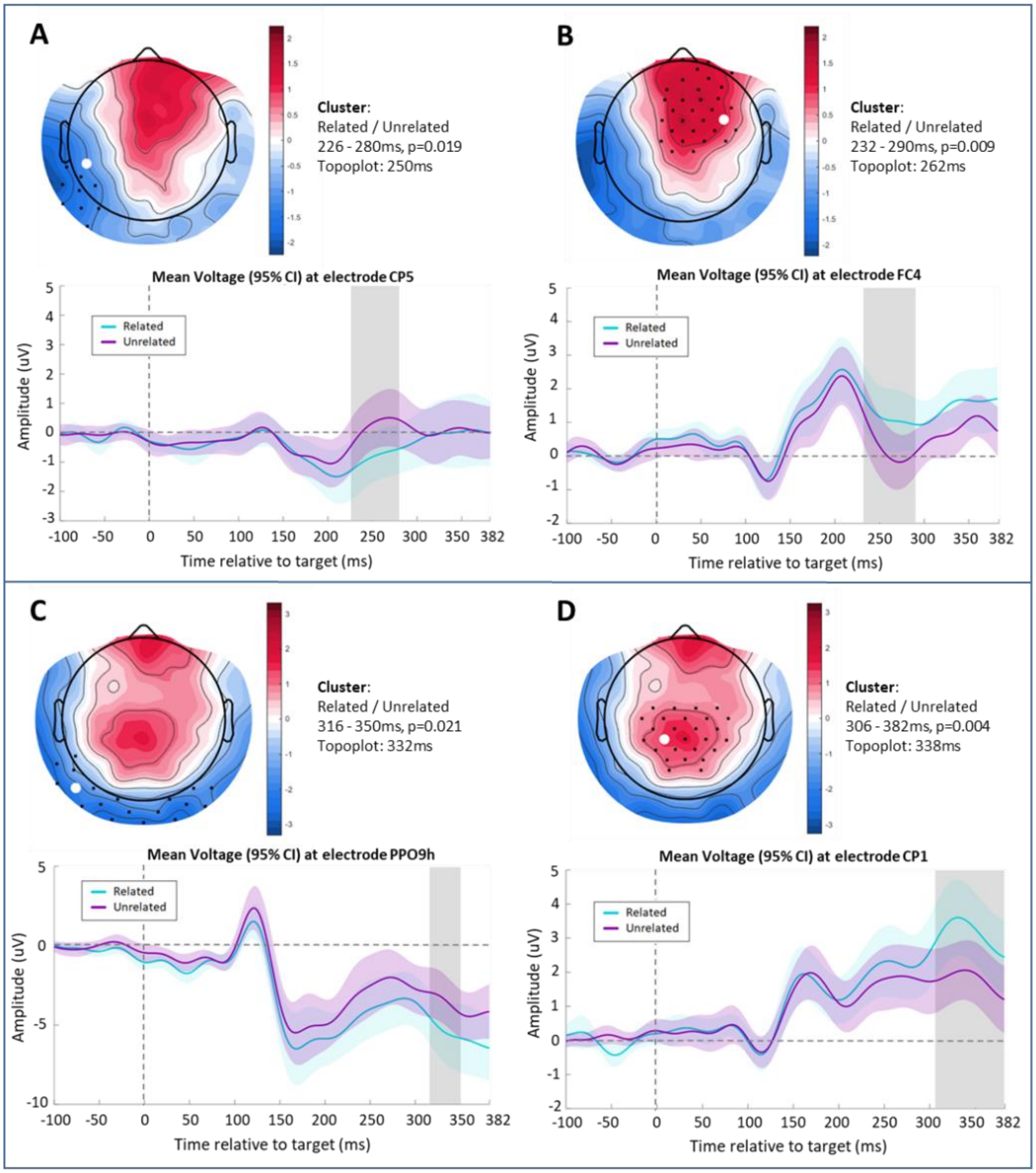


Figure 3.2: *ERP target analysis (related vs unrelated targets). Four main effects from the cluster-based permutation analyses, which contrasted the voltage difference between related and unrelated word-pairs from 0-382ms post-stimulus. ERP scalp topographies revealed two dipolar effects; first, an early fronto-temporal effect at approximately 250ms (A and B); then, a later parieto-occipital effect at around 340ms (C and D). ERP plots show data (mean and shaded 95% confidence interval) from the electrode where the effect was maximal, with the cluster period highlighted in grey.*

As an exploratory analysis, and to increase power to detect a potential interaction effect, we tested for the interaction within each of the main effect clusters by averaging per condition and participant across all channels and time points within each main effect cluster. With this approach, the later negative cluster (C in Figure 3) showed a significant interaction (F (1, 21) = 6.679, p = 0.017, η^2= 0.090), reflecting a larger voltage difference between the related and unrelated targets in a high validity context with respect to a low validity context (other clusters p = 0.396; 0.110; 0.273). Bayesian equivalent analyses considered this to be anecdotal evidence for the alternative hypothesis ($BF_{inclusion}$= 1.505), see Figure 3.3.

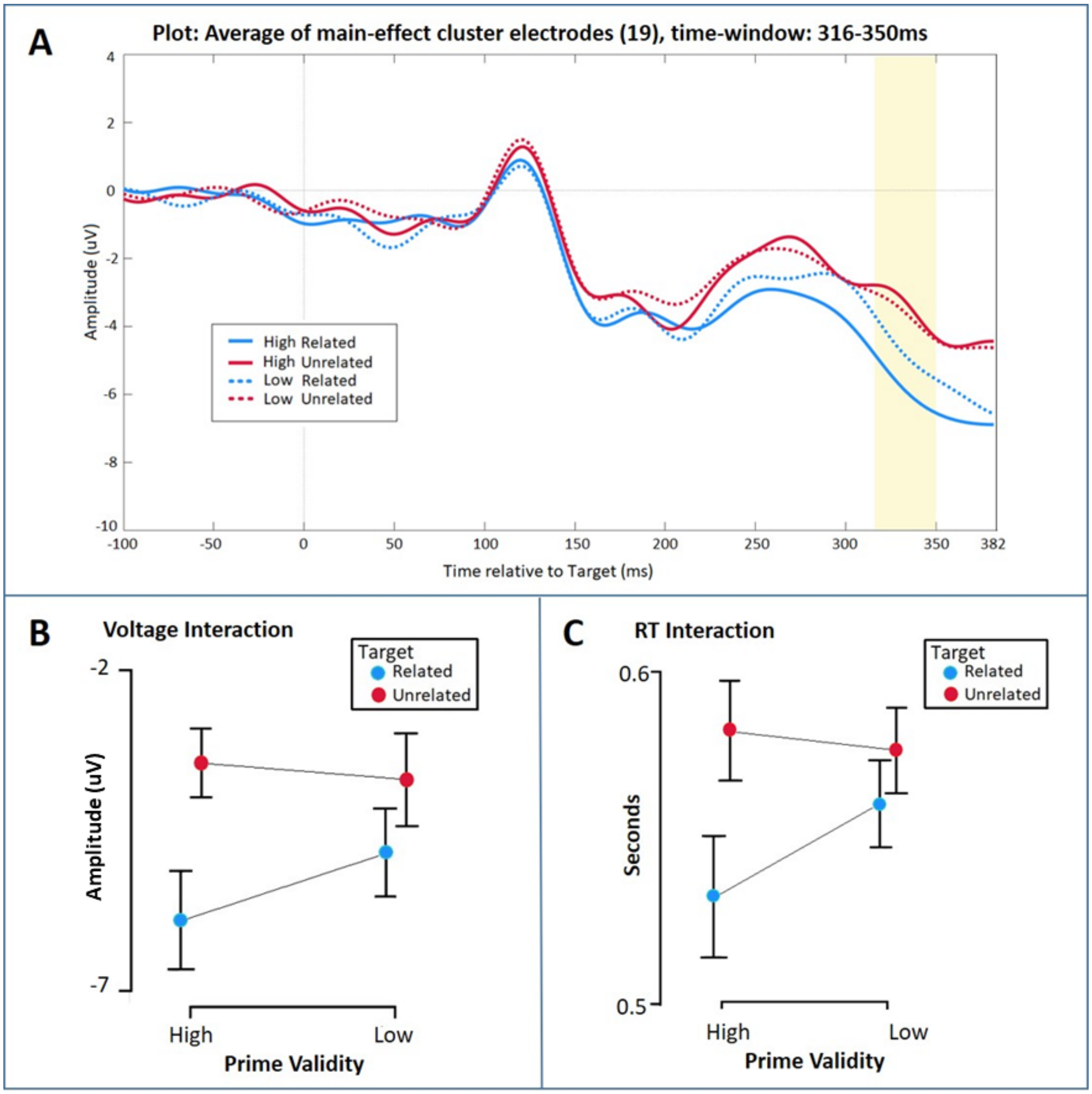


Figure 3.3: *Exploratory ERP analysis to test for the interaction between the four conditions ((HR – HU) – (LR – LU)). The ERP plot in panel A shows the mean of electrodes (19 electrodes) within the 316-350ms cluster found in the main effect analysis (Figure 3, C). Panel B shows the mean for each condition within the same time-window that was analysed with repeated measures ANOVA showing a significant voltage interaction (p = 0.017) with a larger difference in voltage between related and unrelated items in high validity context than low validity context. Panel C shows the significant RTs interaction (p = 0.007) presented in Table 2. In this experiment participant's behaviour (RTs; Panel C) showed the same pattern as their ERP responses (Panel B).*

Source Estimate Analyses

Our pre-registered analyses included whole-brain interaction and main effect contrasts within the time-windows of significant clusters at the sensor level. However, this approach returned no significant clusters at the source level (interaction smallest cluster p = 0.147; main effect smallest cluster p=.067). Furthermore, our preregistered source analyses included regions of interest from the following AAL regions: Left inferior frontal gyrus (LIFG); Left middle temporal gyrus (LMTG); Left superior temporal gyrus (LSTG). However, none of these regions exhibited significant interaction effects or main effects (all FDR corrected p-values > 0.05).

Consequently, for a qualitative visualisation of the source estimates, here we plot the whole-brain thresholded t-values (p<.05) of the source estimate contrasts, uncorrected for multiple comparisons. Specifically, we plot these t-values for the early main effect (Figure 3.3 A and B) and the late main effect (Figure 3C&D) in time windows selected to be entirely within the significant dipolar sensor level clusters (early: 232-280ms; late: 316-350), see Figure 3.4. The thresholded t-values showed the peak of activity at the Right Middle and Superior Fontal Gyri for the early effect; and the activity peak at the Right Supplementary Motor Area for the late effect, as shown in Figure 5.

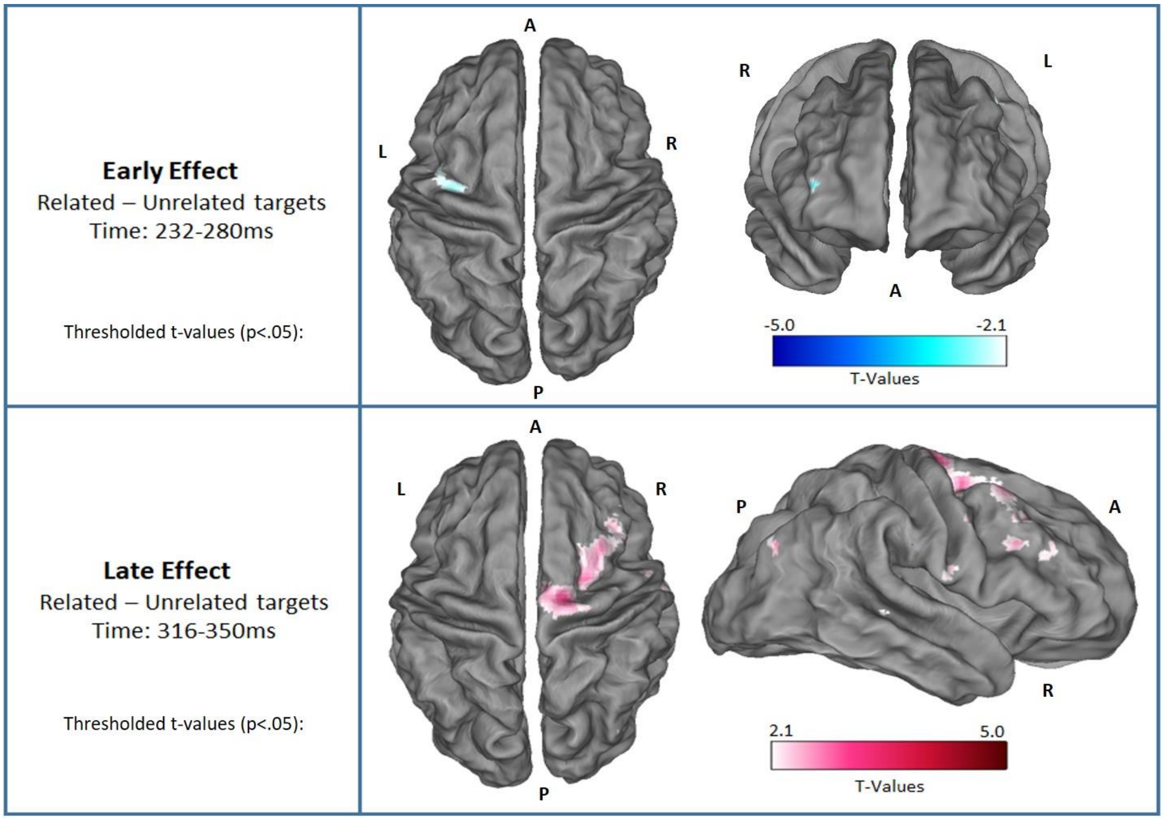


Figure 3.4: *Source estimation analyses. Thresholded t-values (p<.05) of the ERP source estimates over two distinct time windows that corresponded to the early and late ERP effects reported above in Figure 3. In the Figure, the upper panel shows the difference between related and unrelated targets in the early time window (232-280ms), and the lower panel indicates the same difference in a later time window (316-350ms) (thresholded t-values, p < 0.025).*

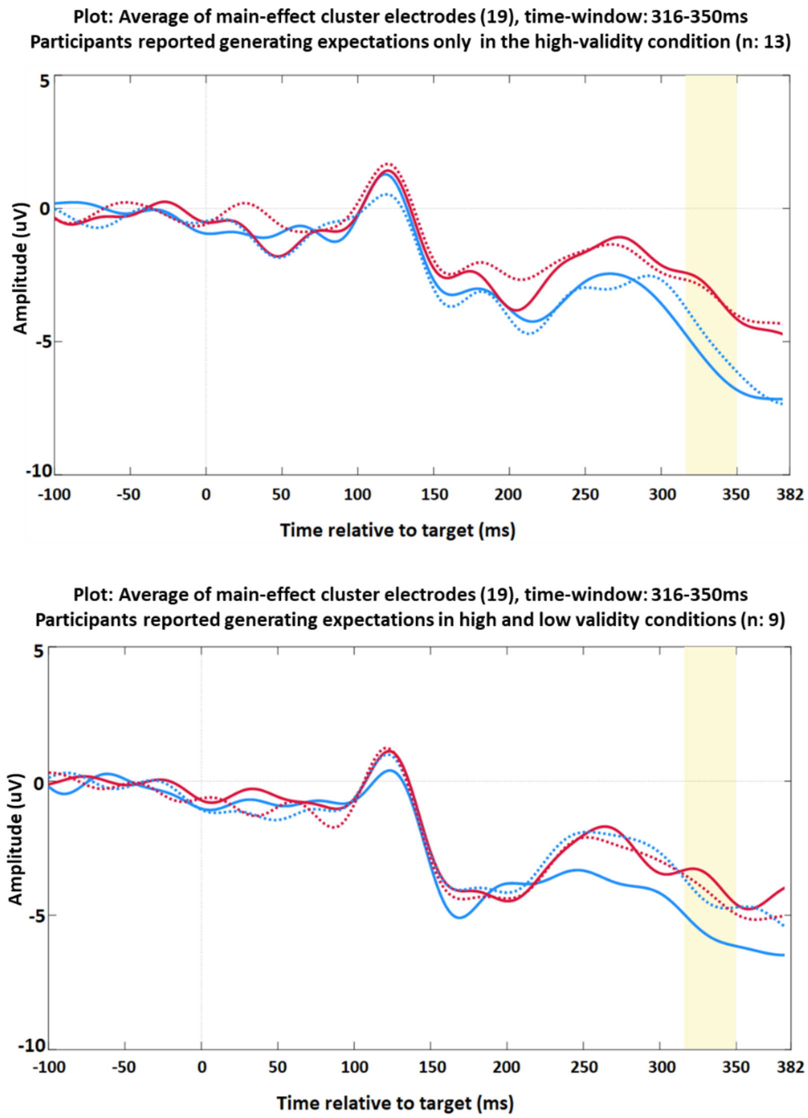


Figure 3.5: *ERP comparison between individuals that reported (self-report form) generating expectations in the high validity condition (upper plot) relative to the ones that reported generating expectations in both conditions (lower plot). We subtracted unrelated from related targets on each validity context and then subtracted both differences ((high related – high unrelated) – (low related – low unrelated)); we next performed an independent samples t-test to compare the size of the effect between both groups that reported either generating expectations in a high validity context, or in both high and low validity context. The results failed to provide evidence for a significant difference between both groups (t = 0.772, p = 0.449, Cohen's d = 0.335).*

Discussion

Predictive coding theory posits that the brain generates expectations about upcoming stimuli at varying levels of complexity – from low-level expectations about stimulus properties through to higher-level conceptual expectations. Here, we investigate the behavioural and electrophysiological correlates of such expectations and their violations at two levels of a semantic expectation hierarchy (local and global). On our two previous behavioural experiments, participants showed evidence of speeded reaction times in related trials relative to unrelated trials, consistent with a local expectation generated about target word identity on the basis of the prime identity. Furthermore, participants generated a more conceptually complex expectation based on the global context (i.e., prime validity) to exhibit greater behavioural facilitation in the high prime validity context than the low prime validity context (Boudewyn et al., 2015). Importantly, in Behavioural Experiment 2 only those individuals who reported conscious strategic expectation showed evidence of behavioural facilitation given by the global context, while those individuals who did not report a conscious strategy only exhibited facilitation as a result of the local context. Together, these behavioural data are consistent with a dissociation between a local expectation about the identity of the target generated by the prime, and a global expectation about the relatedness of the target that necessitates reportable, effortful, and strategic application of expectation.

The present experiment is a replication of the previous Behavioural Experiment 2, and here we include both behavioural and electrophysiological measures, in order to investigate the neural correlates for the use of strategic semantic expectations in this language paradigm. The behavioural results provide evidence for a successful replication of our two previous behavioural studies that are based on the experimental manipulation that was implemented by Hutchison (2007), who also found that the magnitude of the global facilitatory effect was modulated by the level of attentional control (i.e., weaker effect in individuals with lower attentional control; Hutchison, 2007). Our previous Behavioural Experiment 2 suggested that only individuals that reported applying an effortful conscious strategy showed the global context effect as mentioned above; therefore, for this Visual EEG study we only include participants that reported using the prime strategically while performing the task, as the global effect constitutes our effect of interest; and the data of individuals who reported not engaging in strategic processing were not included in the present experiment.

Consistent with this two-stage expectation profile, the ERPs in response to the target words also exhibited a two-stage profile, with an early effect modulated by local expectation (around 250ms) and a later effect modulated by global expectation (around 350ms). These results are broadly consistent with the two-stage profile observed in the auditory oddball local – global paradigm (Bekinschtein et al., 2009), which includes an MMN in an early stage reflecting errors of the local context of the stimuli and a P3b response to errors of the global context given by blocks across the task.

Furthermore, the early effect in the present experiment showed more extreme amplitudes for unexpected targets relative to expected targets, consistent with a prediction error signal, such as the MMN to unexpected/deviant items observed across levels of stimulus awareness (Chennu et al., 2013; Bekinschtein et al., 2009; Faugeras et al., 2012; El Karoui et al., 2015). Moreover, the scalp topography of the early effect has a fronto-central peak, which is consistent with the MMN (Chennu et al., 2013; Bekinschtein et al., 2009; Faugeras et al., 2012), although, its latency is a little longer than seen in some of these previous papers. Additionally, in our source estimation analyses, the early effect was localised to the middle frontal gyrus (Figure 3.4), whereas in another study the local MMN effect was localised to the temporal parietal junction and prefrontal cortex (Chennu et al., 2013), indicating not entirely overlapping neurocognitive processes. Nevertheless, as we observed behavioural semantic priming (as tracked by the early effect) even for participants who were not making strategic expectations, and due to the shared common features with the MMN (i.e., more extreme for errors and with a fronto-central focus), we consider the early effect to be consistent with an error of local expectation – i.e., expectation based on the identity of the prime, rather than the prime validity. Indeed, the MMN is elicited even by individuals who are not actively attending to the stimuli (Bekinschtein et al., 2009).

The late effect, however, was the opposite of what would be expected for a prediction error signal – i.e., its amplitude was more extreme for expected targets compared to unexpected targets. This *apredictive* pattern is not readily explained by prediction error accounts without appeal to precision-weighting, in which a prediction error is weighted by the system's confidence in the signal (Chennu et al., 2013; Friston, 2005). Under precision-weighting, all possible patterns of prediction error signals on the scalp are possible, including apredictive patterns as we observed here, as precision may vary freely across task conditions (Kok, Rahnev, Jehee, Lau, & De Lange, 2012). For example, Barascud, Pearce, Griffiths, Friston, and Chait (2016) reported a larger MEG signal for auditory stimuli that become predictable, relative to stimuli that are entirely unpredictable – i.e., an apredictive pattern – that they linked to up-weighting of the expected stimuli by precision (Heilbron & Chait, 2018). Within predictive coding, attention is one specific mechanism that is thought to increase precision (Hohwy, 2012). Therefore, under a predictive coding framework, one can appeal to varying levels of attention across task conditions. Therefore, we could post-hoc theorise that our late apredictive effect reflects individuals paying greater attention to the high validity trials, as they have a high level of predictability and paying greater attention to related targets than unrelated targets, as the former fulfil their expectations. Therefore, the relative levels of attention across conditions could interact to generate this apredictive effect. Indeed, consistent with this, 59% of our participants (13/22) self-reported that their strategy was to generate an expectation in the high validity condition only (i.e., "I was trying to guess the next word if previous was blue"; where blue was high validity condition); however, when comparing (in figure 3.5) the ERPs from the group of participants that reported generating expectations only in the high-validity condition, relative to participants that reported using both validity contexts to predict the target, there is no interaction effect that provides evidence of a differentiated neural response between both groups. Therefore, it is not clear whether attention directly modulates the ERP signal in this task. Future studies could further examine the varying levels of attention that are required on each condition on a single-subject level, how this effect is reflected on group analyses and how the influence of the attentional resources can affect the neural signal.

An alternative interpretation stems from evaluation of our behavioural data. When comparing the behavioural reaction time interaction with the ERP voltage interaction (see Figure 3.3), both show the same pattern: namely, that the interaction is driven by expected items in a high validity context, showing more extreme values with respect to the other three conditions. This similarity in behaviour and ERP effects suggest that our late 'error' effect may simply reflect processing in service of behaviour, whereby sensory signals are routed to goal-driven analogous motor behaviour (Zylberberg, Slezak, Roelfsema, Dehaene, & Sigman, 2010). Our late apredictive ERP pattern may therefore not reflect a precision-weighted global prediction error, but more simply the result of the brain routing the incoming information into appropriate behaviour. Under this interpretation, our results are therefore also consistent with interpretations of early ERPs as reflections of prediction error and later ERPs as processes related to conscious access and in support of task demands (e.g., Dehaene & Christen, 2011; Rohaut et al., 2015).

It is possible that other later error signals were also evident in the neural response during our task, including those traditionally linked to the N400 (i.e., peaking approximately 400ms post-target). However, we limited our analyses to the 0 to 382ms time-window post-target so as to avoid muscle artefact created by the pronunciation responses. We chose to use a pronunciation task as our aim was to observe the behavioural effect produced by the manipulation of both the local (relatedness) and global context (prime validity) as implemented by Hutchison (2007). Nevertheless, tasks that do not produce large muscular artefacts, such as a lexical decision task (LDT) in which individuals only produce motor responses on filler trials, would allow for analysis of the N400 time-window. However, as argued by Hutchison (2007), participants can complete an LDT with a semantic-matching strategy, meaning that after seeing the target they can verify whether it is related to the prime, which could bias their responses as only words can be related and non-words would be, by their nature, unrelated (Hutchison, 2007). Additionally, as we provided a global context by manipulating the proportion of related items across the task, individuals could bias their responses using the validity cue (Keefe & Neely, 1990); for example, primes that were presented in blue (high validity context) were more likely to be related (80%). Therefore, when seeing a blue prime, individuals could judge their response (word/non-word) solely based on the prime, in this case a 'word' as most of the word-pairs are related. Instead, using a pronunciation task allows for a purer measure of expectations (Hutchison, 2007), with the caveat of limiting the time-window of artefact-free EEG for analysis.

A recent prediction error view on language-related ERPs proposes that the N400 has similar properties to the MMN, as they both are modulated by the predictability of stimuli (i.e., increased ERP amplitude as a prediction-error response) but that their relative latencies indicate prediction-error processing at different levels of stimulus complexity (Bornkessel-Schlesewsky & Schlesewsky, 2019). In our findings, both consecutive effects could be similarly interpreted as reflecting different levels of complexity of precision-weighted prediction error processing across a semantic hierarchy. However, as noted above, appeal to precision-weighting problematically allows for post-hoc explanations of all possible ERP patterns (Bowman, Filetti, Wyble, & Olivers, 2013).

Regarding the source estimation analyses, the early effect was localised to the middle frontal gyrus, which has been previously associated with semantic categorization when compared with passive listening (Noesselt, Shah & Jäncke, 2003). Furthermore, the ERP source estimation analysis for the late effect was localised to the supplementary motor area, consistent with the above interpretation that the late interaction reflects goal-driven routing toward action. Indeed, this area has been linked to speech motor control, verbal working memory, and predictive top-down mechanisms in speech perception (Hertrich, Dietrich, & Ackermann, 2016). However, neither of these two regions were part of our pre-registered hypotheses. Therefore, these source estimates should be interpreted with caution, and future studies with this paradigm will wish to replicate these sources.

In our pre-registered analyses, we also hypothesised that we would observe electrophysiological markers of differential expectations generated by the high and low validity primes, prior to the onset of the target. Specifically, we expected these differential expectations to be reflected in the ERPs, including the slope of a putative slow wave (Chennu et al., 2013), and/or in the power of the EEG in the alpha/beta bands, as these have been previously associated with the precision of expectations (Bauer, Stenner, Friston & Dolan, 2014). However, we found no evidence of any differences in these measures between high and low validity primes prior to target onset. One interpretation is that our specific measures were simply not sensitive enough to detect the differential expectations in these conditions. Indeed, we powered our study to detect the post-target behavioural effect specifically. An alternative interpretation is that expectations were, in fact, not different between the two conditions. Indeed, under predictive coding, the brain is considered to optimize the difference between its expectations and sensory input by updating its internal model (Friston, 2010); hence, it is possible that the optimal means of minimising prediction error in this task is to always predict the related target, regardless of the prime validity. For example, even if one were to consciously expect that an upcoming target will be unrelated (as in a low validity trial), it is simply not possible to accurately predict the identity of that target, as the range of possible unrelated target words is considerable. Therefore, even though predicting the identity of a specific related target had only a ~22% probability of being correct in a low validity context, it was still more likely than predicting any one of the vast arrays of potential unrelated target words. Future inspection of participants' meta-cognition in relation to their specific expectations following prime presentation will help speak to this interpretation.

In conclusion, we here reported ERP evidence of hierarchical matching of semantic expectations to incoming speech. Lower lever expectations based on the local context (i.e., the prime identity) elicited an early and predictive pattern that matches with prediction error accounts. Higher level expectations generated from the global context required awareness of the global rule and the use of a reportable strategy and were associated with an apredictive pattern that can be interpreted within a precision-weighted prediction error account or may reflect the routing of sensory signals and their expectations into task-directed behaviour.

CHAPTER 4: ELECTROPHYSIOLOGICAL MARKERS OF AUDITORY STRATEGIC SEMANTIC EXPECTATIONS

Introduction

This thesis has so far explored the relatedness proportion paradigm in a semantic priming task as implemented by Hutchison (2007), aiming to measure the use of strategic semantic expectations when individuals process language stimuli. The RP paradigm enables the measurement of trial-by-trial semantic local expectations by presenting individuals with related or unrelated word-pairs. Simultaneously, semantic global expectations can also be measured on a trial-by-trial basis by providing individuals with a global context across the task. The global context involves two different contexts, where each have either a higher or a lower probability of containing related word-pairs, and this probability is given to participants as a cue incorporated in the prime (e.g., coloured primes). As we previously proposed, this paradigm follows the same structure as the local-global paradigm (Bekinschtein et al., 2009) by measuring expectation violations at shorter (local) time scales, that are generated from bottom-up processing; and the violation of expectations at longer (global) time scales that require the recruitment of top-down processing, as global expectations involve the context in which stimuli are presented (Bekinschtein et al., 2009).

On a local level, our behavioural results from three previous behavioural experiments show that individuals present priming effects, indicating faster responses when pronouncing related targets relative to unrelated targets. Semantic priming effects are largely reported in the literature (Hutchison, 2007; Neely, 1976; Neely, 1991; Lau et al., 2013; Klinger et al., 2000; Chwilla et al., 1998, etc.); and faster responses for related targets are a result of facilitation given by expectancy (Hutchison, 2007; Becker, 1980) or automatic mechanisms (Collins & Loftus, 1975; Quillian, 1967). Automatic semantic priming mechanisms are recruited when the prime and target are presented with a short gap in between (e.g., 200ms SOA) (Neely & Keefe, 1989); whereas a long gap between prime and target presentation (e.g., 1240ms) allows for the use of expectancy mechanisms, where a specific set of expectations about the target are generated from the prime (Hutchison, 2007). The experiments exposed in the previous chapters include long SOA manipulations (i.e., 800 and 1240) providing the conditions for expectancy mechanisms to operate (Hutchison, 2007); although these priming effects can be considered as less automatic than priming effects in experiments using shorter SOA manipulations, we cannot assert that these effects rely fully on consciously controlled mechanisms either.

However, on a global level we provide evidence for the use of strategic semantic expectations in this paradigm. In our three previous behavioural experiments, individuals show significant RPEs, meaning larger priming effects in high validity contexts relative to low validity contexts (Hutchison, 2007; de Groot, 1984; Neely et al., 1989; Keefe & Neely, 1990). These results reflect facilitation under a highly valid context in which is highly likely that the target will be related to the prime (high prime validity context), so individuals use this contextual information to generate global expectations about the target, therefore showing faster responses for trials that fulfil these expectations (related targets) (Hutchison, 2007; Keefe & Neely, 1990; Kuperberg & Jaeger, 2016). Furthermore, evidence for the use of strategy while performing this task, stems from individuals' self-report about the use of strategy in our previous experiments, as only individuals who report being strategic (i.e., using the prime validity to generate global expectations about the target) show behavioural global effects (RPEs) relative to non-strategic individuals where only a behavioural local effect is detected (priming effects).

In the previous chapter, following the behavioural effects mentioned above, we explore the neural correlates for the use of semantic strategic (global) and non-strategic (local) expectations to process language stimuli in this paradigm. The results of our previous visual EEG experiment show an early ERP effect around 250ms at fronto-temporal electrodes, showing a predictive pattern where the ERP signal shows more extreme values for unexpected (unrelated) than expected (related) items, which goes in line with previous predictive coding accounts where the ERP signals reflect the prediction error response (Dehaene & Christen, 2011; Rohaut et al., 2015). Furthermore, we report a later ERP effect around 350ms at parieto-occipital electrodes with an apredictive pattern that can be either a result of precision-weighted prediction error (Kok et al., 2012) or represent the routing of sensory signals that are oriented to task-direct behaviour due to task demands (e.g., pronunciation task) (Zylberberg et al., 2010). The later effect also yielded a follow-up voltage interaction in which the ERP signal shows a greater difference for related and unrelated targets (priming effects) under a high validity context, relative to a low validity context (RPE), consistent with our previous behavioural results and with previous RP manipulations that have reported this ERP interaction (Lau et al., 2013; Brown, Hagoort & Chwilla, 2000; Holcomb, 1988). An interesting finding is that the voltage interaction follows the same pattern as behavioural responses giving rise to the explanation for sensory signals being routed to meet task demands (Zylberberg et al., 2010), which was previously mentioned.

These ERP results resemble the local-global paradigm findings in healthy participants, where a two-profile error detection signal is observed consistently across studies; including an early prediction error effect (MMN) followed by a later positivity (P3b) that reflects the global error as it interacts with the global context (Bekinschtein et al., 2009; Faugeras et al. 2012; King et al., 2014; El Karoui et al., 2015; Chennu et al., 2013). Moreover, both errors show a predictive pattern as the amplitude of the ERP signal is more extreme for unexpected items in both local and global contexts (Bekinschtein et al., 2009); however, our previous ERP results reflect the error in the early 'local' ERP effect, but not in the 'global' ERP effect where there is an apredictive ERP signal as mentioned above. There are differences in the results as both paradigms represent distinct experimental manipulations, where the local-global paradigm (Bekinschtein et al., 2009) uses auditory tones and usually instruct participants to count global deviant occurrences, whereas this specific RP manipulation (Hutchison, 2007) employs word-pairs that are presented visually and instruct participants to pronounce the target (motor task) in each trial. However, both tasks follow the same underlying structure managing to differentiate the use of expectations at a local context that involves automatic or at least non-strategic processing, from global expectations that are built from the use of the global context across time, involving consciously controlled strategic processing (Bekinschtein et al., 2009; Hutchison, 2007). Therefore, these 'local' and 'global' common features are reflected in the ERP two-stage profile observed in both paradigms, and its detection in this RP language paradigm complies with the overall aim of this thesis, which is to identify the use of strategic semantic expectations when processing language stimuli.

Consequently, so far, we have achieved the aim of identifying a paradigm and neural markers of strategic semantic expectations. The present chapter builds on these previous results with the aim to adapt this experimental manipulation so it can be tested in patients diagnosed with disorders of consciousness, that may still preserve residual language comprehension abilities. In order to implement a clinically viable paradigm for DOC patients, who remain unresponsive, we aim at lowering task demands to adjust to patients' level of processing and thus obtain a more accurate estimate of their abilities. Patients diagnosed with DOC present difficulties fixating their eye movements towards relevant stimuli (Ting, Perez Velazquez, & Cusimano, 2014), where successful visual fixation and pursuing objects with their eyes can be observed in MCS patients and is considered as a sign of consciousness recovery (Royal School of Physicians, 2020). As a consequence of the brain injury, several functional systems within the brain are disrupted, hindering eye movements coordination and fixation (Ting et al., 2014). Therefore, stimuli are usually presented auditorily to patients, and evoked potentials (time-lock EEG responses) are used to measure auditory pathways functionality (Harrison & Connolly, 2013). For example, in a study by Coleman and colleagues (2009) all DOC patients showed auditory cortex activation in response to sound irrespective of their behavioural diagnosis, suggesting auditory stimuli as a good choice to assess signs of awareness in DOC patients. Hence, the next task will present word-pairs auditorily to both healthy participants and DOC patients, where the global context (high and low prime validity) will be given by presenting the primes with distinctive type of voices (male and female) to cue participants about the probability of the target being related (i.e., 77.8% related targets in a high prime validity context and 22.2% related targets in a low prime validity context).

Moreover, our previous studies include a pronunciation task that aimed at detecting a behavioural effect that could provide evidence for the use of strategic semantic processing on a trial-by-trial basis; however, this motor task is not suited for patients' abilities and it will not be included in the present experiment. As the pronunciation task allowed us to test for the presence of underlying behavioural effects, reflecting the use of both local (priming effects) and global (RPE) semantic expectations, as a compensation we are including some aspects that would allow us to check for the use of strategic expectations. First, we are instructing participants to do a mental task in which they have to guess the target whenever they hear the prime. Several studies have reported DOC patients following mental tasks when they preserve the cognitive abilities to do so (Cruse et al., 2012; Monti et al., 2010; Gibson et al., 2014; Holler et al., 2013; Bodien, Giacino, & Edlow, 2017; Horki et al., 2014). Asking individuals to guess the targets (think ahead) intends to encourage them to use expectations on a trial by trial basis. In fact, several individuals' self-report from our previous studies reported guessing what the next word is when they described their strategy use (e.g., "I would try to guess the related word if it was blue, in order to say it faster", or "I would think of likely candidates for the second word"). Moreover, we are giving participants the same instructions as previous experiments, by showing them the global rule that explains the type of voice that correspond to a specific context (e.g., "if the prime is said by a female voice the word-pair is highly likely to be associated”). Therefore, individuals can follow the mental task and at the same time using the global rule across the task. In addition, as we encourage participants to predict the targets and follow the rule, we are also including the strategy self-report form for healthy participants as previously used (Vidal-Gran, Sokoliuk, Bowman, & Cruse, 2020; Chapter 2 and 3 of the present thesis), to check whether participants report using a strategy while performing the task, although this self-report form is not included for patients.

The RP manipulation presented in this thesis measures expectations based on both local and global contexts on a trial-by-trial basis, meaning that a trial can be locally related or unrelated (local context), but belongs simultaneously to a globally high or low prime validity as it is pronounced by either a male or a female voice (coloured primes in previous manipulations). The same occurs in the local-global paradigm, where one trial is locally deviant or standard, and simultaneously globally deviant or standard. However, the global regularity is built across the block in which the trial is presented, and this regularity can be inferred after encountering several trials (Bekinschtein et al., 2009); whereas in this RP manipulation (Hutchison, 2007) the global regularity is defined by the prime cue (type of voice), thus a global expectation can be formed only by the presence of a single trial. Because DOC patients have regular fluctuations in their level of arousal and attentional resources (Giacino et al., 2014), evaluating the use of strategic processing (global expectations) in a trial-by-trial modality is more appropriate to meet their abilities, since a global context given across a block (such as in the local-global paradigm; Bekinschtein et al., 2009), requires patients to keep their attentional resources on task throughout the duration of the block. However, in this RP manipulation using global expectations in a single trial would still require that the individual understood the global rule and knows that they need to apply it to a specific trial, which could arguably involve similar levels of attentional resources as the local-global approach.

The following hypotheses are stated in our pre-registration form (link in methods below), which were established a priori before start collecting the data and were based on our previous EEG experiment results (Vidal-Gran, Sokoliuk, Bowman, & Cruse, 2020; Chapter 2 and 3 of the present thesis). First, we hypothesised that on a local level individuals would show greater ERP amplitudes for unexpected targets in contrast to expected targets (prediction error signal) in an early stage around 250ms post-target, if they are using the prime to predict the target. The signal reduction for related items would reflect semantic facilitation given by the predictive mechanisms, and the signal increase for unrelated items would suggest the detection of the prediction error signal (Lau et al., 2013; Kuperberg & Jaeger, 2016). On a global level, if participants are using the voice cue (prime validity cue) that represent the global context across the task, we will detect a greater difference between the ERP amplitudes of related and unrelated targets that are presented in a high prime validity context, relative to a low prime validity context. Moreover, as it was observed in our previous Visual EEG study we expect to observe an apredictive signal in this later effect (i.e., more extreme voltage values for expected targets relative to unexpected targets), contrary to previous findings that have reported more extreme values for unexpected targets in contrast to expected targets (Lau et al., 2013; Brown et al., 2000).

This global effect will be evident at a later stage in time with respect to the early local effect mentioned above.

In some trials we would ask healthy participants to rate their confidence on their guesses (more details are provided in the methods section). If participants are using the global context to engage in strategic expectations, they will have higher confidence ratings for the word-pairs presented in a high validity context than in a low validity context. Furthermore, if participants are using the global context to expect upcoming semantic stimuli, we will observe an ERP global main effect in which both global standard conditions will differ from global deviant conditions, more details will be provided in the methods section where we establish a direct comparison with the conditions from the local-global paradigm (Bekinschtein et al., 2009). Moreover, in our visual EEG experiment (Vidal-Gran, Sokoliuk, Bowman, & Cruse, 2020; Chapter 3 of the present thesis) there was no evidence for a differentiated processing of primes in high validity relative to low validity contexts as it was expected, and the slow wave before the target presentation did not show differences in prime validity contexts either. Evidence suggest that the generation of expectations in different conditions can be reflected in the ERP signal (Chennu et al., 2013) or in the power of the EEG in alpha/beta bands (Bauer et al., 2014). Therefore, in the present experiment we will test for these prime differences, even though we did not observe them in our previous EEG study Vidal-Gran, Sokoliuk, Bowman, & Cruse, 2020; Chapter 3 of the present thesis).

EEG Auditory Experiment

Methods

The present study was pre-registered in the Open Science Framework website, where details can be found in the following link: https://osf.io/z5pma. Deviations from the pre-registered methods and analyses will be appropriately reported throughout the text.

Healthy Participants

We recruited participants through the Research Participation Scheme website of the University of Birmingham, and participants chose to receive either credits or £10 per hour of participation. We recruited 42 participants, and we discarded the data of two participants as there was an EEG malfunction at the time of recording. The sample was composed of 40 participants (median age: 21, range: 18 – 35). We defined this sample size prior to data collection, where we performed two non-parametric power calculations, one including behavioural data and the other using ERP data. For the behavioural calculation, we pooled the reaction times from Behavioural Experiment 2 and EEG Visual Experiment (reported in chapter 2 and 3, respectively) resulting in data from 84 participants and we used replacement. Both calculations were performed in a Matlab script (Mathworks, Inc., Natick, Massachusetts) to estimate the power to detect the effect using a two-way repeated measures ANOVA, where alpha (α) is defined at 0.05. Next, we select with replacement N participants and performed the same two-way repeated measures ANOVA 1000 times to test for each interaction effect (behavioural and ERP), and p-value of the two-way interaction is stored in each itineration. The power value at N corresponds to the proportion of stored p-values which fall below alpha. According to the calculation, an estimated sample size of 32 participants was necessary to achieve 81% of statistical power to detect an interaction at $p<0.05$. Regarding the ERP power calculation, we included the ERP data from the EEG Visual experiment (reported in Chapter 3), where we found a voltage interaction around 350ms, so we selected the time-points from the significant cluster of the main effect (related vs unrelated) and created an average per participant for each condition. We include data from 22 participants and used replacement using the same procedure mentioned above. The results yielded an estimated sample size of 28 participants to achieve 80.5% of statistical power to detect a significant difference in a repeated measures t-test at $p<0.05$.

Even though the power calculations yielded sample sizes of 32 participants for behavioural data and 28 participants for ERP data, the current task design has significant differences from previous experiments, such as the change from visual to auditory stimuli and the absence of a motor task (i.e., pronunciation). Therefore, we define a sample size of 40 healthy participants to increase statistical power, which yielded 99% of statistical power to detect the same behavioural interaction effect at $p<0.05$; and 91% power to detect the same voltage interaction at $p<0.05$.

Individuals reported being mono-lingual native English speakers; right-handed; not having previous history of neurological conditions or diagnosis of dyslexia; and no hearing impairment. This study was approved by the STEM Ethical Review Committee of the University of Birmingham, England. All participants gave written informed consent prior to their participation in the study.

Stimuli

We used the same stimuli as in the EEG Visual Experiment; however, the word-pairs were now presented auditorily instead of visually. The prime validity is presented to participants using type of voices (either male or female voices) instead of colours as in the previous experiments (either yellow or blue). To create the auditory word-pairs, we synthesised speech with VOICEBOX speech processing toolbox (Brookes, 1997) for Matlab (Mathworks, Inc., Natick, Massachusetts), using both a male and a female voice. There was a total of 704 words including prime and target words, and each word was created for each type of voice (1408 spoken words). The female voice had a median length of 444ms (range: 205ms - 951ms), and the male voice a median length of 442ms (range: 154ms - 1069ms). We did not find evidence that would indicate that the lengths of the words of both types of voices are different ($p=0.824$). Additionally, the Bayesian t-test found strong evidence in favour of the null hypothesis ($BF_{10} = 0.043$), which stated that there are no significant differences between the lengths of the male and female voices.

The relatedness of a prime-target pair was defined as the local context in this design, where a related prime-target pair represents a local standard trial (locally expected); otherwise, an unrelated prime-target pair was accounted as a local deviant trial (locally unexpected). Furthermore, we provided a global context by presenting high and low validity primes in either a male or female voice, which cue participants about the targets' identity (i.e., likelihood of target being related or unrelated to the prime). Related prime-target pairs in a high validity context, and unrelated prime-target pairs in a low validity context are considered global standard trials (globally expected); whereas unrelated prime-target pairs in a high validity context, and related prime-target pairs in a low validity context represent global deviant trials (globally unexpected), see table 4.1.

Prime Validity	Relatedness	Local context	Global context
High	Related	Standard (LS)	Standard (GS)
High	Unrelated	Deviant (LD)	Deviant (GD)
Low	Related	Standard (LS)	Deviant (GD)
Low	Unrelated	Deviant (LD)	Standard (GS)

Table 4.1*: Experimental conditions of the present experiment (columns 1 and 2) compared with conditions from the local-global paradigm by Bekinschtein et al. (2009) in columns 3 and 4.*

Procedure

Stimuli were displayed to participants using Psychtoolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) in Matlab (Mathworks, Inc., Natick, Massachusetts). The task had two types of trials; some trials were designed for ERP data analyses, and others to analyse confidence ratings; both type of trials are shown in Figure 4.1. In the ERP trials (332 trials), participants were requested to guess the upcoming word (target) as soon as they hear the prime, and no motor response was required. Each trial started with a black fixation cross on a grey background for 750ms; next, the prime word was played through the headphones in either a male or a female voice, while the fixation cross stayed on the screen. Subsequently, 1240ms (SOA) after the prime onset, the target word was played in the same type of voice as the prime; whilst the fixation cross continued to be displayed on the screen, until 1300ms after target onset. Next, the screen remains blank for 1000ms. For the confidence rating trials (20 trials), we instructed participants to write down their responses in a blank table that we provided them. Each trial started with a black fixation cross in a grey background for 750ms, followed by the presentation of the prime through the headphones in either a male or female voice. Next, the instructions appeared on the screen instructing participants to write down on the first column of the answer sheet the prime word they just heard and then pressing any key to continue. The next instructions indicated participants to write down in the second column of the answer sheet their guess about the upcoming target and then press any key to continue. Subsequently, another instruction was displayed on the screen instructing participants to indicate in the third column of the answer sheet their level of confidence for their target guess response. The confidence rating scale ranged from 1 to 4, where each number had the following label: 1) Not confident at all; 2) Slightly confident; 3) Fairly confident; 4) Completely confident. Next, the target word was played through the headphones while the fixation cross was on the screen; followed by the instruction to write down on the fourth column of the answer sheet, the target word they just heard. The EEG data of the confidence rating trials were no later subjected to statistical analyses; however, the confidence ratings were used to test statistical differences between ratings in a high validity context (10 trials) and a low validity context (10 trials). As the ERP trials did not include any active task (i.e., button press, pronunciation, etc.), the confidence rating trials also served the purpose of keeping participant’s attention throughout the task, as these were presented randomly across the experiment.

Each participant was tested individually and seated approximately 70cm away from the computer screen. All participants received written information about the study, the instructions and signed the consent form to participate in the study. We instructed participants that they will listen on the headphones the first word of the word-pair (prime) in either a male or a female voice. Moreover, to encourage the use of predictions we told them that as soon as they hear the prime, they had to try to guess which word was likely to be heard next as the second word of the pair (target), and that they would hear the actual target after a time gap. Participants were told that the type of voice of the prime will give a clue about the probability of the target being associated or not associated to the prime. Half of participants received the following written instructions: “If the first word is said by a female voice, it is highly likely that both words will be associated with one another; and if the first word is said by a male voice, it is highly likely that both words will not be associated with one another” (Hutchison, 2007). The other half of participants received the opposite instructions, as they were assigned with the inverse voices. All participants completed four practice trials with the experimenter’s supervision.

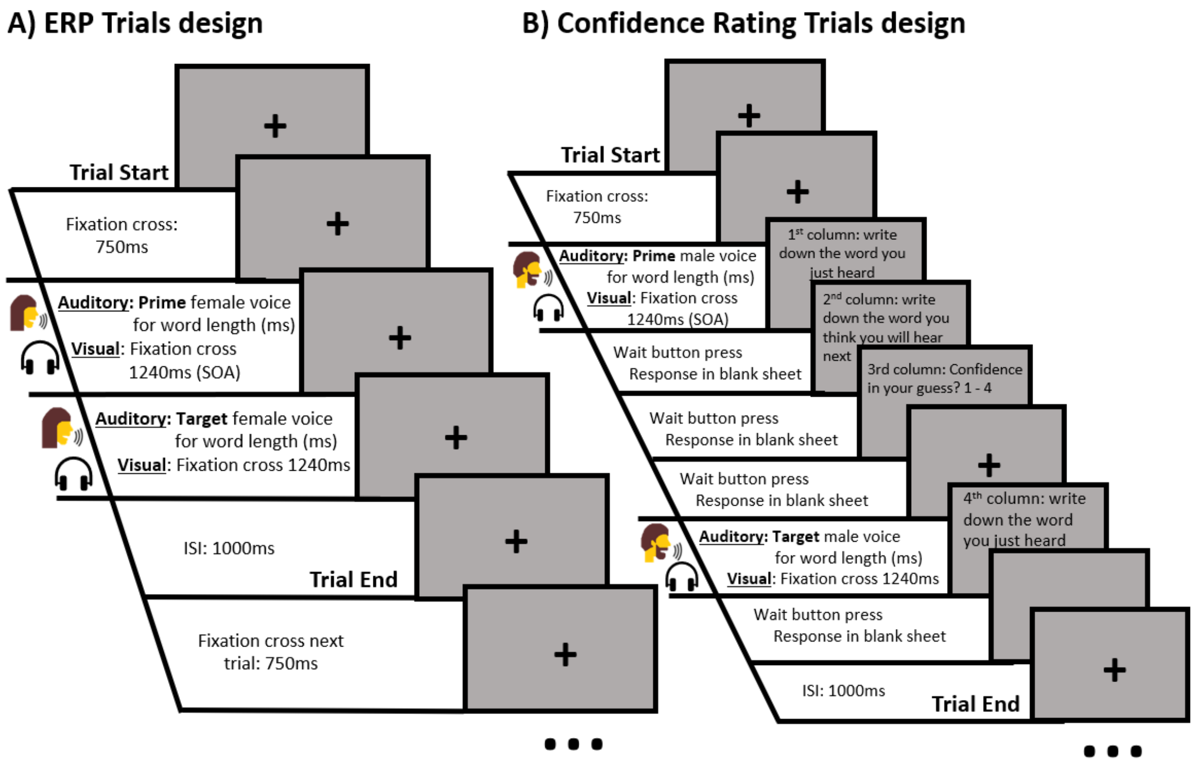


Figure 4.1: *Semantic Priming Relatedness Proportion task (Hutchison, 2007) where the prime validity manipulation was done using either a male or a female voice in each experimental trial. Panel A shows the trial procedure for the type of trials where the ERP responses were considered for the statistical analysis. Panel B shows the type of trials that measured confidence rating about the target guess.*

To evaluate whether participants were using strategic expectations while doing the task, we asked them to complete a self-report form at the end of the experimental procedure. The self-report form included the following questions: 1) Which voice was highly likely to be associated? (Responses: MALE / FEMALE); 2) Did you use the type of voice as a cue for knowing whether the following word was associated or not associated? (Responses: YES / NO); 3) Describe your strategy on each trial.

<u>EEG recording</u>

We continuously recorded the EEG signal of all participants with a high-density 125 channel AntNeuro EEG system (AntNeuro b.v., Enschede, Netherlands). The brain activity was acquired at a sampling rate of 500Hz, and we kept impedances below 20 kΩ when possible. We placed the ground electrode on the skin at the left side of the forehead, and we previously sanitised the skin at this area and applied conductive gel. The voltage that was measured by the 127 electrodes on the shielded waveguard cap was referenced to the CPz channel. Moreover, we recorded ECG activity using external sensors. We placed two ECG electrodes, one on the left side of the chest and the other electrode on the right side.

<u>EEG Pre-Processing Pipeline</u>

The automated EEG pre-processing pipeline is identical to the one reported in EEG Experiment 1 (Chapter 3). Once we run the automated script in all participants, we plot all trials for each participant to manually check that this procedure eliminated noisy trials and/or channels accurately. For participants that keep noisy activity after the automated procedure, we perform manual pre-processing instead, where we reject trials and/or channels manually, next we run ICA and manually remove components to eliminate artefacts. We then interpolate the rejected channels, apply baseline correction (-200) and use average referencing. If any participant's data keeps showing significant noise (with respect to other participants) after the automated and manual pre-processing, we will discard the data and replace it with a new participant's data. The data of 2 participants presented significant noise after both the automated and manual procedures, as a consequence of an EEG equipment defect at the time of recording, therefore we discarded their data and replace it with 2 new participants. From a total of 40 participants, the automated EEG pre-processing script was successful in 29 participants, whilst 11 participants required manual EEG pre-processing following the procedure mentioned above.

Analyses

Confidence ratings analyses

In the confidence rating trials, participants were requested to rate their confidence in their guess about the target-word based on the prime-word, which included responses ranging from 1 to 4. Each participant had 20 trials, which were randomly presented across the experiment; moreover, half of these trials (10) were presented in a high validity context and the other half (10) in a low validity context, and the context was cued to participant with the type of voice in which the word-pair was played. Of a total of 800 confidence rating trials across participants for both contexts, 77 trials (9.6% of all trials) were removed from the analysis as individuals did not provide a confidence response (6 trials); they misunderstood the prime word (69 trials); or they provided more than one confidence rating (2 trials). The remaining ratings from all 40 participants were included in the analyses, where we estimated the median of each condition from each participant for the statistical analysis. Thus, a median of 9 trials (range: 6-10) contributed to the high validity condition; and a median of 9 trials (range: 7-10) contributed to the low validity condition.

We investigated whether individuals were using the global context when guessing the most likely target; therefore, we contrasted their confidence ratings when word-pairs were presented in a high validity context and a low validity context. If participants were using strategic expectations that rely on the global context, we expected higher confidence ratings in a high validity context with respect to a low validity context. We conducted the analyses in JASP 0.9.2.0 software (JASP Team, 2019) by performing a paired samples Wilcoxon signed-rank test; moreover, we conducted a Bayesian paired sample t-test.

To test whether participants show higher confidence ratings when making related guesses relative to unrelated guesses under a low validity context, we conduct a follow-up analysis where we divide participants into two groups, according to the type of response they give in low validity trials, either related or unrelated. Participants that showed mixed responses (e.g. three related and seven unrelated guesses in low validity trials) were excluded from this analysis. In total, 5 participants' confidence ratings were excluded as they give mixed responses and 35 participants responded either related (n: 15) or unrelated (n: 20) in low validity trials. In order to balance the groups, we only considered the first 15 participants of the unrelated group resulting in 15 participants per group. We conducted an

Independent samples Welch test, and a Bayesian Independent samples t-test using the median responses from a low validity context. The group of participants with Low-Related responses included 15 participants (trials median: 9; range: 7 - 10) and the group of Low-Unrelated responses included 15 participants (trials median: 9; range: 8- 10).

EEG Analysis:

Target ERP, Prime ERP and Prime time frequency analyses:

We analysed the time-courses (ERPs/time-frequency) within several time-window of interest. For the analyses of the primes, we selected a time-window from prime presentation up to target presentation (0-1240ms). Regarding the target analyses, the choice of time-windows was based in the results of EEG Experiment 1, in which we found an early ERP effect around 250ms and a late ERP effect around 350ms. The ERP analyses in the present study are similar to EEG experiment 1; thus, we defined the target time-windows from 100-500ms for the early effect (related/unrelated contrast); and 300 to 800ms for the late effect (related/unrelated – high validity/low validity interaction). The time-windows were compared with the cluster mass method of the open-source Matlab toolbox FieldTrip (Oostenveld et al., 2011), which was described in EEG Experiment 1.

Prime slow wave linear fit analyses:

Similar to EEG experiment 1, we intended to further test for ERP evidence of expectation formation in response to the prime; therefore, we analysed whether a slow wave differentiates high validity and low validity conditions. For this comparison we used a least-squares linear fit to the averaged ERPs of each condition (High and Low validity primes) for each electrode and participant (as per Chennu et al., 2013). Next, the slope values were compared between conditions with the spatial cluster mass analysis in FieldTrip.

Results

The Wilcoxon signed-rank test showed that participants were more confident in their guesses about the upcoming target when the primes were presented in a high validity context, in contrast when the primes were in a low validity context ($W = 703$, $p < 0.001$, $r = 0.715$). Median values for high validity and low validity primes were 3 and 1, respectively (see figure 4.2). Moreover, the Bayesian Paired Sampled t-test showed extreme evidence in support of the alternative hypothesis (BF_{10}= 2.362e+13), which states that primes in a high validity context have higher confidence ratings in contrast with primes in a low validity context.

An independent samples Welch t-test failed to provide evidence that participants who responded with related guesses under a low validity context were more confident in their responses in contrast to participants that responded with unrelated guesses ($t = 1.694$, $p = 0.101$, Cohen's $d = 0.618$). In addition, the Bayesian independent samples t-test yielded anecdotal evidence in support of the null hypothesis, which claims that there is no difference in the confidence ratings of participants that responded related guesses in a low validity context, than participants that gave unrelated responses in the same context (BF_{10}= 0.996).

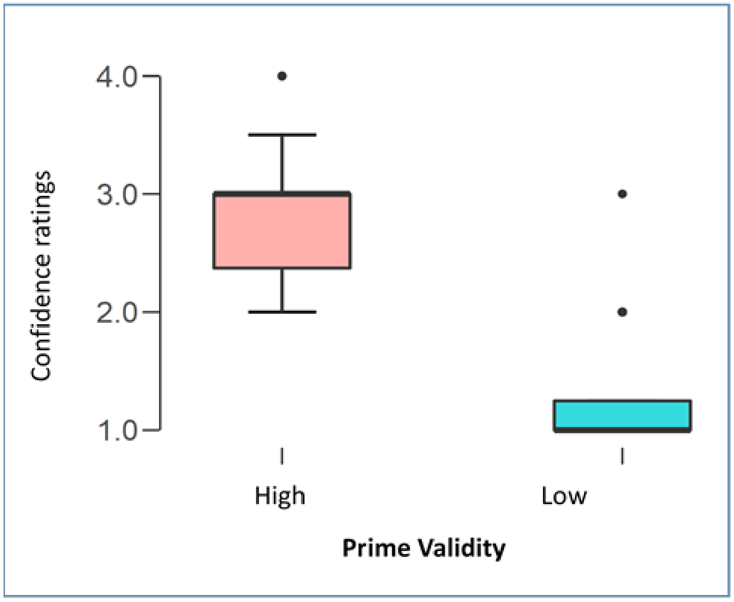


***Figure 4.2**: Tukey Boxplot showing confidence ratings' median values across participants (high validity vs low validity), the bold black line indicates the sample median and dots indicate outliers; the colour bars indicate the interquartile range (IQR). Confidence ratings indicate that participants were more confident in their guesses when the primes were in a high validity context, relative to a low validity context.*

EEG Results

Prime analyses: ERPs, time frequency and slow wave linear fit analyses

As our pre-registered analyses stated, we first analysed the prime main effect (high validity – low validity) in a time-window from prime onset up to 1240ms. In the ERP analysis there was no expectation effect in response to the prime, as primes in a high validity context did not significantly differ from primes in a low validity context (smallest cluster $p = 0.05$; see figure 4.3). Furthermore, the slow wave linear fit analysis did not show any difference between both conditions as no clusters were formed. As an exploratory follow-up slow wave linear fit analysis, we defined a shorter time-window (248 to 1240ms) based on the peak of the global field power index; and the analysis showed that no clusters were formed either. Regarding the alpha (8-12Hz) and beta (13-30Hz) time frequency analyses, there were no significant clusters in either the alpha band (smallest cluster $p = 0.12$) or the beta band where no clusters were formed.

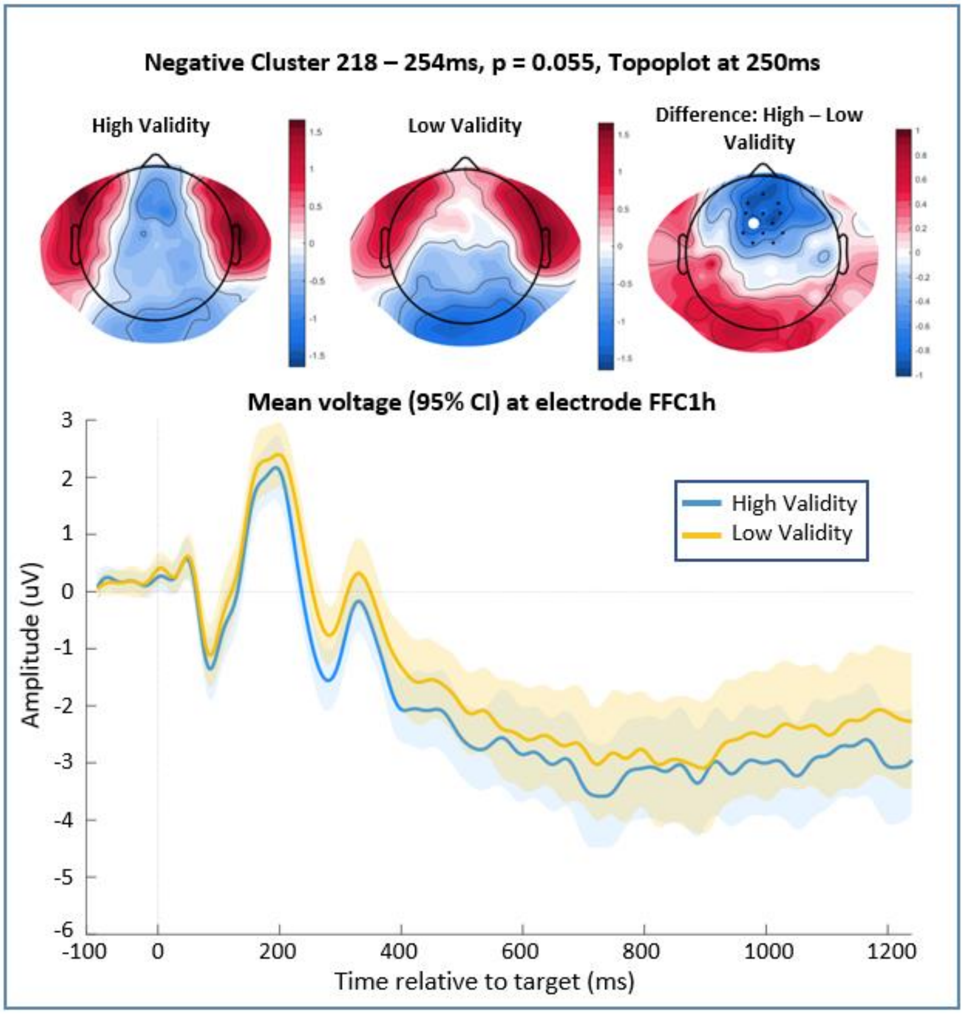


Figure 4.3: *ERP prime analysis (high vs low validity). ERPs containing Mean voltage (95% CI) at electrode FFC1h, contrast between high and low validity primes at a time window from 0 to 1240ms post prime and pre-target, revealed no significant difference between conditions (smallest cluster p = 0.055).*

Target results: ERPs

The pre-registered interaction contrast (High Validity/Low Validity, Related/Unrelated) was conducted in a time window from 300 to 800ms; and the cluster-based permutation analysis indicated no significant interaction between the relatedness of the target with the validity of the prime (smallest cluster p = 0.08), see figure 4.4.

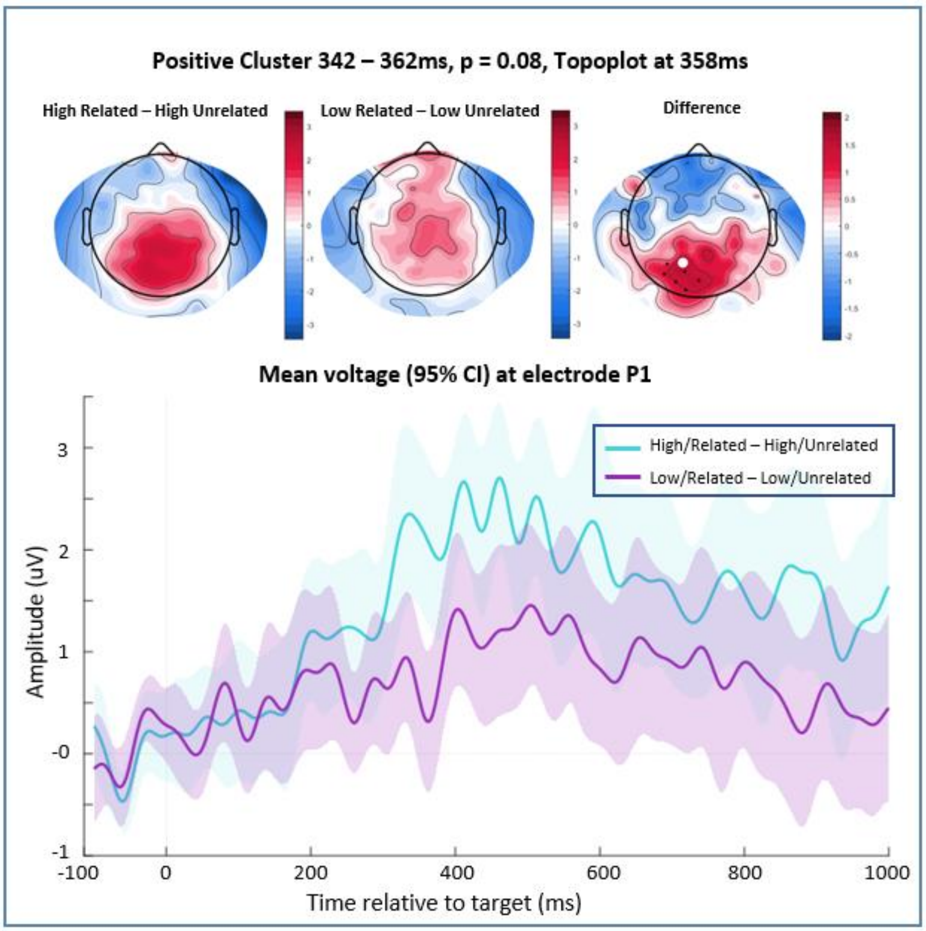


Figure 4.4: *ERP interaction analysis (prime validity*relatedness target). Mean voltage (95% CI) at electrode P1, shows the difference between related and unrelated targets in a high validity context, relative to the same difference in a low validity context. Time window from 300-800ms post target, failed to detect a significant difference between conditions (smallest cluster p = 0.08).*

As stated in our pre-registered analyses, we tested for a main effect of relatedness (related vs unrelated targets) in an early time window from 100 to 500ms. The cluster-based permutation analysis showed a parietal electrodes main effect of relatedness of the target at approximately 350ms post-stimulus (positive cluster: 172 – 500ms, p = 0.001), see figure 4.5. The voltage in the cluster showed that unrelated targets had more extreme values with respect to related targets; therefore, the effect reflects a predictive signal as it was expected from our pre-registered analyses.

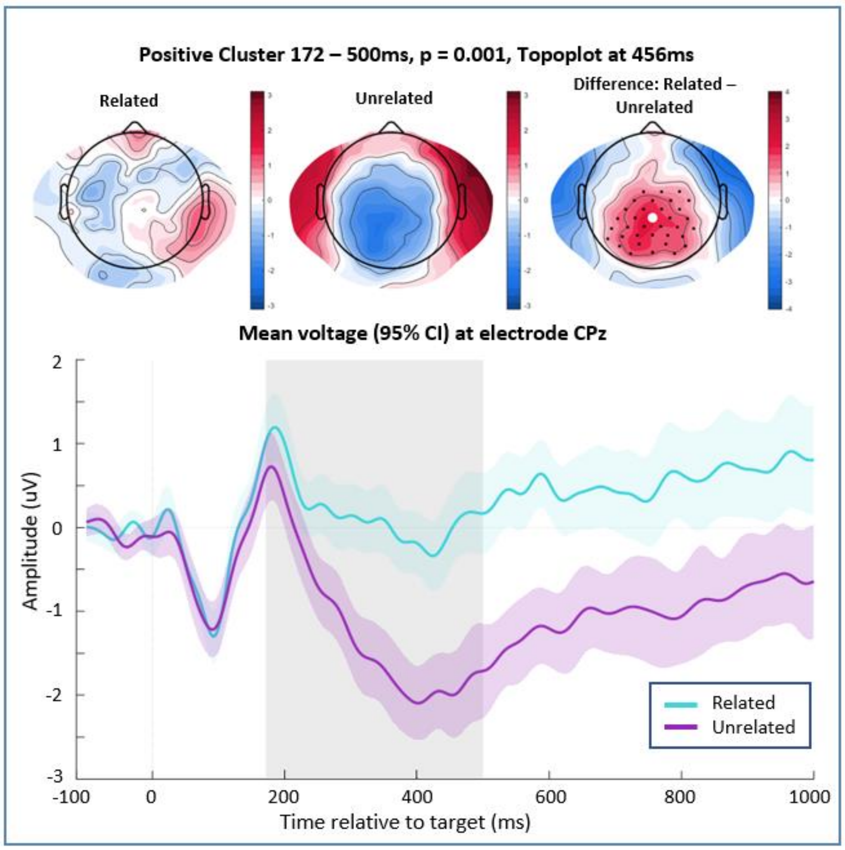


Figure 4.5: *ERP target main effect analysis (related vs unrelated). ERPs containing Mean voltage (95% CI) at electrode CPz, which shows the difference between related and unrelated targets in a time window from 100-500ms post target. The analyses yielded a significant difference between conditions (smallest cluster p = 0.001).*

Moreover, as our pre-registered follow-up analyses indicated we ran an average interaction in the later time window from 300-800ms, by averaging per condition and participants across all channels and time points within the main effect cluster, which was identified at parietal electrodes ($p = 0.001$), in Figure 4.6. According to the results, we found a significant follow-up interaction between the relatedness of the target with the validity of the prime ($F(1, 39) = 7.636$, $p = 0.009$, $\eta^2 = 0.164$); which showed a larger voltage difference between the related and unrelated targets in a high validity context in contrast to a low validity context. Furthermore, the Bayesian Repeated Measures ANOVA showed substantial evidence in favour of this interaction ($BF_{inclusion} = 3.118$), see figure 4.7.

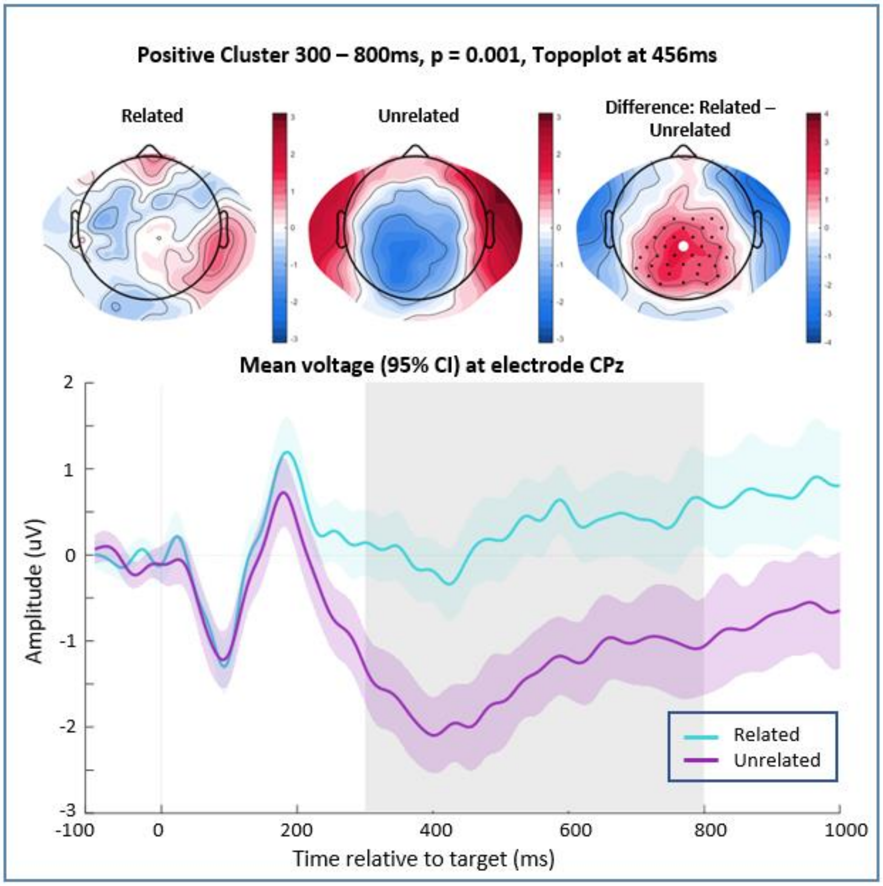


Figure 4.6: *ERP target main effect analysis (related vs unrelated)**.*** *ERPs containing Mean voltage (95% CI) at electrode CPz, which shows the difference between related and unrelated targets in a time window from 300-800ms post target. The analyses yielded a significant difference between conditions (smallest cluster p = 0.001).*

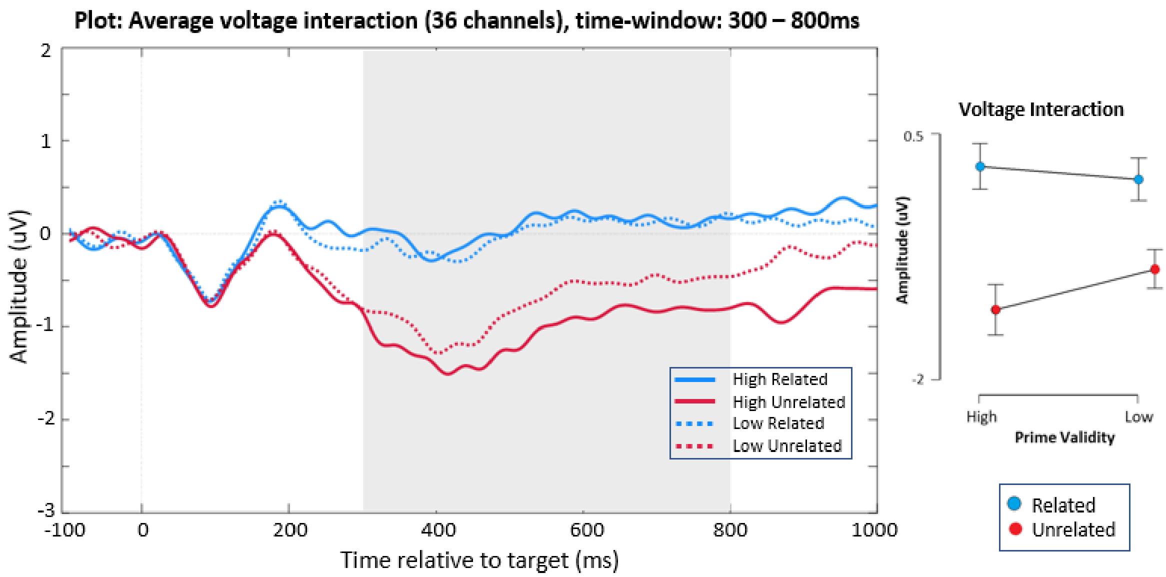


Figure 4.7: *ERP average voltage interaction (validity of the prime * relatedness of the target). The ERP plot shows the mean of electrodes (36 electrodes) within the cluster from 300 – 800ms. The mean for each condition within the same time-window show a significant voltage interaction (p = 0.009) with a larger difference in voltage between related and unrelated items in high validity context relative to a low validity context.*

As a post-hoc exploratory analysis, we ran the same voltage average interaction as above but in a previous time-window from 100-300ms, to check whether there is evidence for the interaction prior to the pre-registered later time window (300-800ms). For this analysis we ran the interaction in the positive cluster shown on Figure 4.5, although we only include the cluster values from 100ms to 300ms. The results failed to provide evidence for an interaction between the relatedness of the target and the validity of the prime in a time window from 100-300ms ($F(1,39) = 2.532$, $p = 0.12$, $\eta^2 = 0.061$). The results yielded a significant main effect for relatedness ($F(1,39) = 54.888$, $p < 0.001$, $\eta^2 = 0.585$), showing that related targets significantly differ from unrelated targets. Moreover, a Bayesian Repeated Measures ANOVA showed anecdotal evidence in favour of the interaction null hypothesis, which states that there is no evidence for an interaction between the relatedness of the target and the validity of the prime ($BF_{inclusion} = 0.575$). In addition, this analysis provided extreme evidence for the main effect alternative hypothesis, which indicates a significant difference between related and unrelated targets ($BF_{inclusion} = 9.201e+7$).

As indicated in our pre-registered analyses, we tested for a main effect between global standard and global deviant trials in a time-window from 300 to 800ms and the cluster based permutation analysis failed to detect a significant main effect of global context (smallest cluster $p = 0.089$), see figure 4.8.

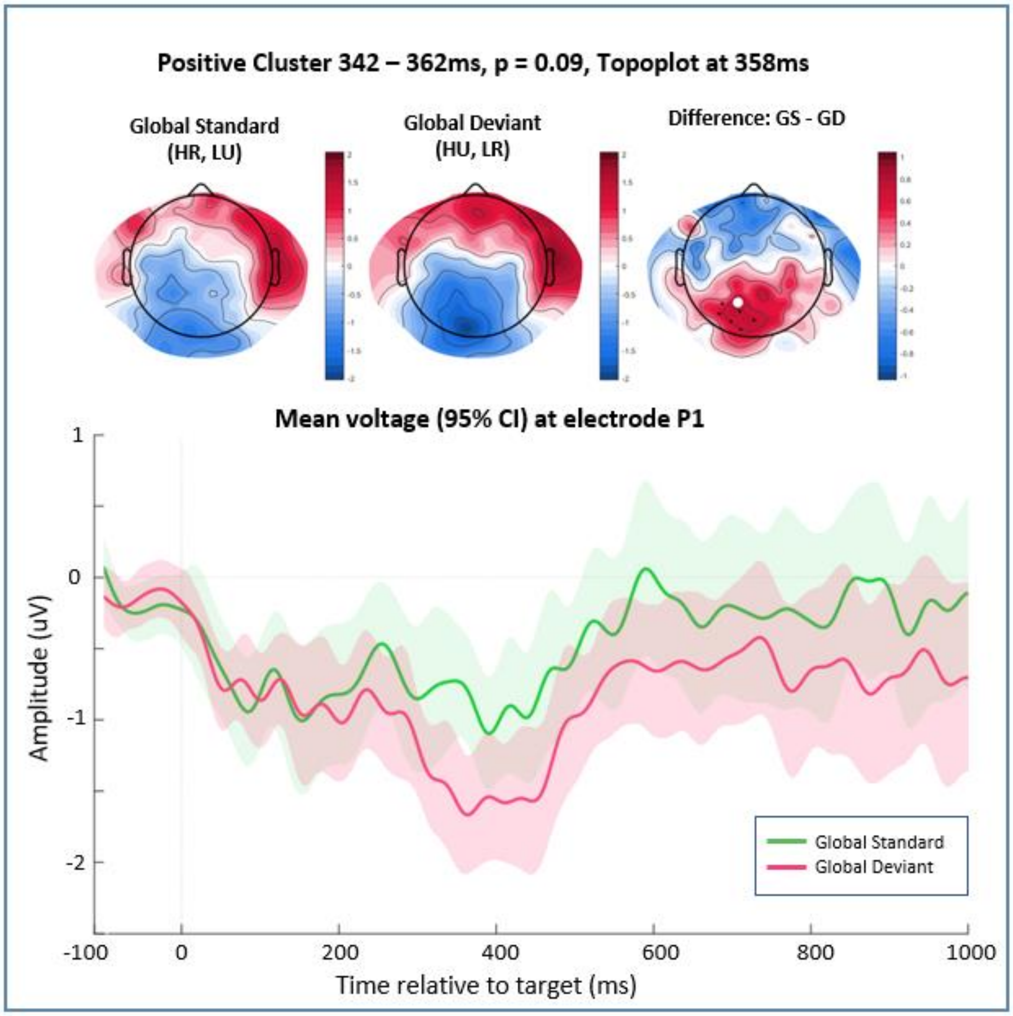
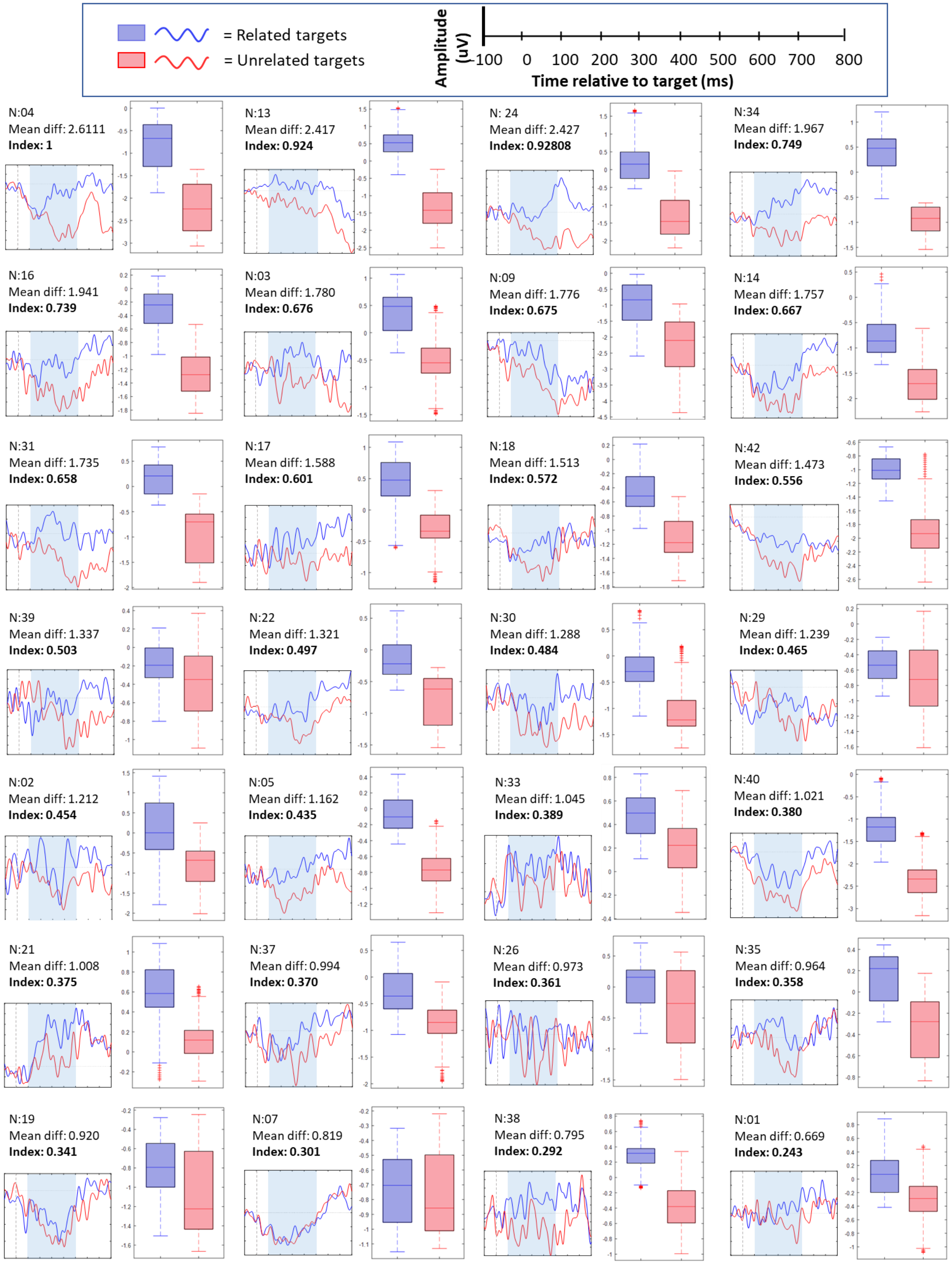


Figure 4.8: *ERP target main effect analysis (Global standard vs Global deviant). ERPs containing Mean voltage (95% CI) at electrode P1, which shows the difference between global standard and global deviant trials in a time window from 300-800ms post target. According to the analyses there was not a significant difference between global conditions (smallest cluster p =* 0.089*).*

As exploratory follow-up analyses, we include single subject analyses for both the local and global effects. For the local effect we conduct a t-test on each participant's earlier time window from 100-500ms, by averaging per condition and trials across all channels within the main effect cluster shown in figure 4.5. The results yielded that 37 out of 40 (94,9%) participants showed significant differences between the voltage values in the related and unrelated conditions (all p values < 0.001) and 3 participants failed to show evidence for this difference (N: 7, $p = 0.075$; N: 10, $p = 0.333$; N= 32, $p = 0.149$), see table 4.1 for all t-test results. Moreover, for visualisation purposes we subtracted each participant's cluster condition mean (Related - Unrelated) and applied min-max normalization to each participant's difference in order to scale single-subject local effects from the largest (1) to the smallest effect (0); therefore, single-subject local effects in figure 4.9 follow a scaled (index) order.

Regarding the global effect, we ran a follow-up average interaction in the later time window from 300-800ms on each participant's data as indicated by the main effect cluster displayed in figure 4.6. Using each participant's dataset, we averaged the cluster voltage values of each trial on each condition (high related, high unrelated, low related and low unrelated) and the resulting time points were subjected to an analysis of variance ANOVA independent samples. The results failed to show evidence of an interaction between the validity of the prime and the relatedness of the target at a single subject level in all participants (see table 4.2 for all ANOVA results), as it was observed on a group level depicted figure 4.7, see figure 4.10 for each participant's cluster interaction.


= Related targets
= Unrelated targets
Amplitude (uV)
-100 0 100 200 300 400 500 600 700 800
Time relative to target (ms)
N:04
Mean diff: 2.6111
Index: 1
N:13
Mean diff: 2.417
Index: 0.924
N: 24
Mean diff: 2.427
Index: 0.92808
N:34
Mean diff: 1.967
Index: 0.749
N:16
Mean diff: 1.941
Index: 0.739
N:03
Mean diff: 1.780
Index: 0.676
N:09
Mean diff: 1.776
Index: 0.675
N:14
Mean diff: 1.757
Index: 0.667
N:31
Mean diff: 1.735
Index: 0.658
N:17
Mean diff: 1.588
Index: 0.601
N:18
Mean diff: 1.513
Index: 0.572
N:42
Mean diff: 1.473
Index: 0.556
N:39
Mean diff: 1.337
Index: 0.503
N:22
Mean diff: 1.321
Index: 0.497
N:30
Mean diff: 1.288
Index: 0.484
N:29
Mean diff: 1.239
Index: 0.465
N:02
Mean diff: 1.212
Index: 0.454
N:05
Mean diff: 1.162
Index: 0.435
N:33
Mean diff: 1.045
Index: 0.389
N:40
Mean diff: 1.021
Index: 0.380
N:21
Mean diff: 1.008
Index: 0.375
N:37
Mean diff: 0.994
Index: 0.370
N:26
Mean diff: 0.973
Index: 0.361
N:35
Mean diff: 0.964
Index: 0.358
N:19
Mean diff: 0.920
Index: 0.341
N:07
Mean diff: 0.819
Index: 0.301
N:38
Mean diff: 0.795
Index: 0.292
N:01
Mean diff: 0.669
Index: 0.243

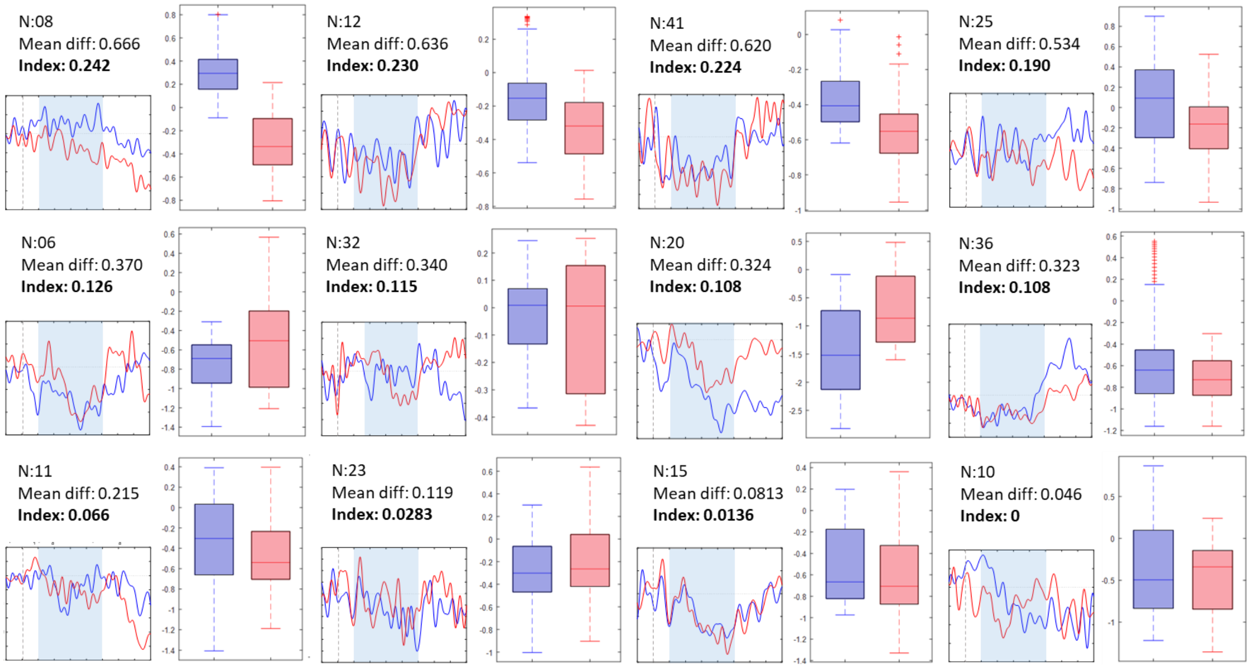


Figure 4.9: *Single-subject local effect. Each participant includes a participant number (N); the cluster mean difference between Related and Unrelated targets (Mean diff); and the min-max normalization value (index) that orders participants from the largest (1) to the smallest (0) local effect. The plot on the left of each participant shows an average of the cluster channels for the related (blue) and unrelated (red) conditions, where the presence or absence of an N400 effect can be observed; the light blue rectangle indicates the time window (100-500ms) for the local effect and the amplitude values (y axis) were omitted as they are the same as the adjacent boxplot. The boxplot on the right indicates the cluster voltage for the related (blue) and unrelated (red) conditions; the central line shows the median; the bottom and top edges refer to the 25th and 75th percentiles; and outliers are represented with a +.*

Participant	p value	Corrected p		CI min	CI max	Related Cluster Mean	Unrelated Cluster Mean	Mean Difference (R-U)	Rescaled difference (index)
4	<0.001	<0.001	*	1.286	1.561	-0.560	-3.171	2.611	1.000
24	<0.001	<0.001	*	1.395	1.656	0.123	-2.304	2.427	0.928
13	<0.001	<0.001	*	1.840	2.001	-0.292	-2.709	2.417	0.924
34	<0.001	<0.001	*	1.301	1.454	0.693	-1.274	1.967	0.749
16	<0.001	<0.001	*	0.876	1.003	-0.166	-2.107	1.941	0.739
3	<0.001	<0.001	*	0.798	0.965	0.578	-1.203	1.780	0.676
9	<0.001	<0.001	*	1.233	1.392	-1.236	-3.012	1.776	0.674
14	<0.001	<0.001	*	0.877	1.011	-1.207	-2.965	1.757	0.667
31	<0.001	<0.001	*	1.026	1.165	0.041	-1.694	1.735	0.658
17	<0.001	<0.001	*	0.692	0.837	0.633	-0.955	1.588	0.601
18	<0.001	<0.001	*	0.587	0.722	0.006	-1.506	1.513	0.572
42	<0.001	<0.001	*	0.852	0.962	-1.167	-2.640	1.473	0.556
39	<0.001	<0.001	*	0.132	0.268	0.559	-0.778	1.337	0.503
22	<0.001	<0.001	*	0.590	0.735	0.005	-1.316	1.321	0.497
30	<0.001	<0.001	*	0.679	0.852	0.136	-1.152	1.288	0.484
29	<0.001	<0.001	*	0.107	0.258	-0.398	-1.638	1.239	0.465
2	<0.001	<0.001	*	0.764	1.000	0.201	-1.010	1.212	0.454
5	<0.001	<0.001	*	0.668	0.729	0.129	-1.034	1.162	0.435
33	<0.001	<0.001	*	0.246	0.321	0.549	-0.496	1.045	0.389
40	<0.001	<0.001	*	1.098	1.232	-1.342	-2.363	1.021	0.380
21	<0.001	<0.001	*	0.399	0.471	1.117	0.109	1.008	0.375
37	<0.001	<0.001	*	0.522	0.676	-0.064	-1.058	0.994	0.370
26	<0.001	<0.001	*	0.290	0.514	0.098	-0.874	0.972	0.361
35	<0.001	<0.001	*	0.430	0.515	0.434	-0.530	0.964	0.358
19	<0.001	<0.001	*	0.241	0.320	-0.649	-1.569	0.920	0.341
7	0.021	0.075		0.005	0.059	-0.735	-1.554	0.819	0.301
38	<0.001	<0.001	*	0.598	0.680	0.559	-0.236	0.795	0.292
1	<0.001	<0.001	*	0.261	0.396	0.047	-0.622	0.669	0.243
8	<0.001	<0.001	*	0.570	0.665	0.166	-0.500	0.666	0.242
12	<0.001	<0.001	*	0.161	0.213	-0.187	-0.822	0.636	0.230
41	<0.001	<0.001	*	0.164	0.226	-0.481	-1.101	0.620	0.224
25	<0.001	<0.001	*	0.168	0.318	-0.124	-0.658	0.534	0.190
6	<0.001	<0.001	*	-0.278	-0.191	-0.873	-1.243	0.370	0.126
32	0.042	0.149		0.001	0.078	-0.079	-0.419	0.340	0.115
20	<0.001	<0.001	*	-0.861	-0.751	-1.798	-1.475	0.324	0.108
36	<0.001	<0.001	*	0.089	0.199	-0.536	-0.859	0.323	0.108
11	0.004	0.015	*	0.041	0.211	-0.333	-0.117	0.215	0.066
23	0.001	0.002	*	-0.141	-0.040	-0.366	-0.485	0.119	0.028
15	<0.001	<0.001	*	0.040	0.094	-0.954	-0.872	0.081	0.014
10	0.097	0.333		-0.017	0.200	-0.770	-0.724	0.046	0.000

Table 4.2: *Single-subject local effects. The table shows the results of t-tests on each participant's time window from 100-500ms, by averaging across trials and channels within the main effect cluster showed in figure 4.5, the t-test contrasts between related and unrelated targets. The table includes the p-values, corrected p-values, confidence intervals, related cluster mean, unrelated cluster mean, the difference cluster means (related-unrelated), and the rescaled difference as calculated with a mix-max normalization that orders effects from the largest (1) to the smallest (0).*

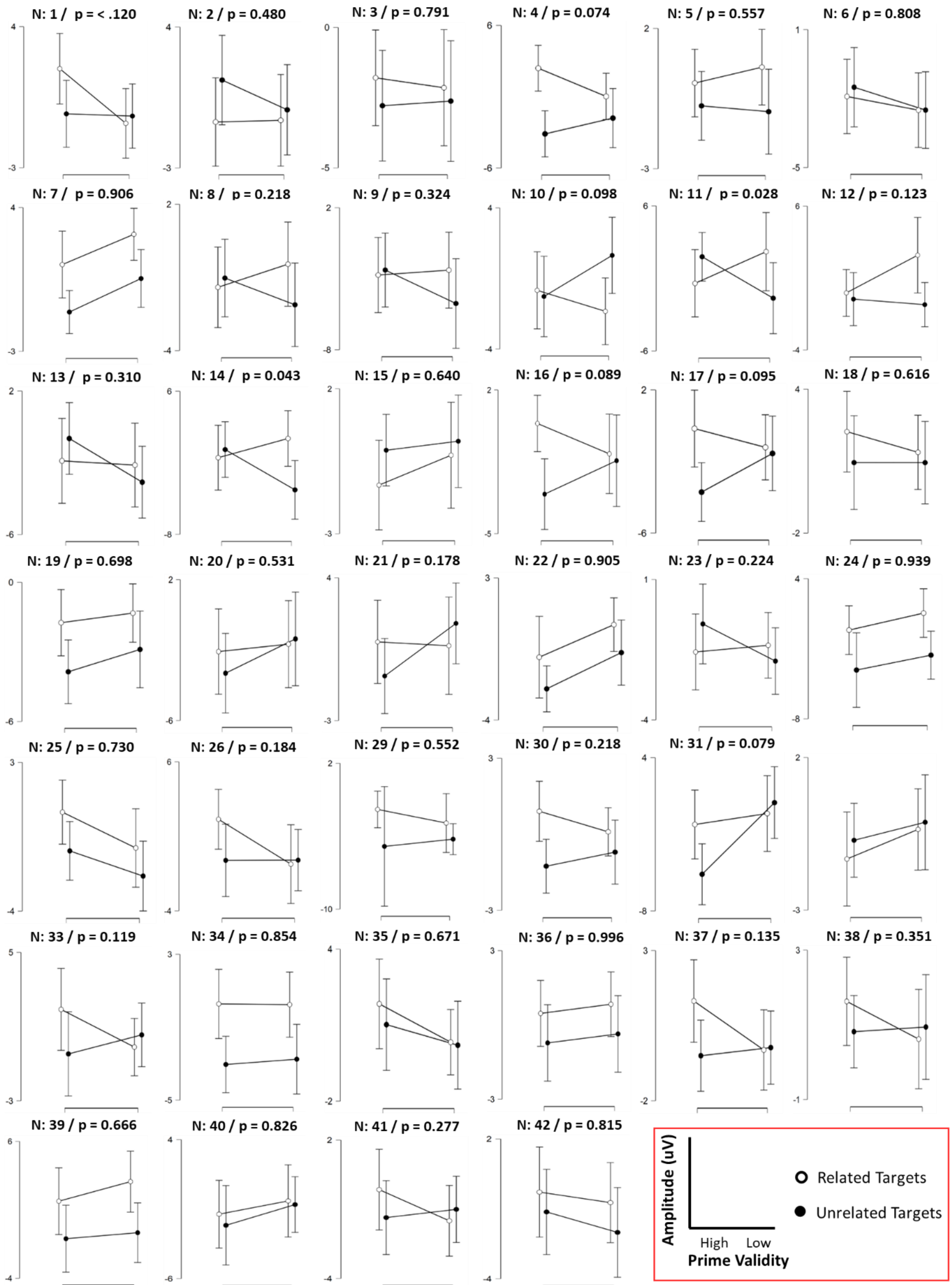

N: 1 / p = < .120
N: 2 / p = 0.480
N: 3 / p = 0.791
N: 4 / p = 0.074
N: 5 / p = 0.557
N: 6 / p = 0.808
N: 7 / p = 0.906
N: 8 / p = 0.218
N: 9 / p = 0.324
N: 10 / p = 0.098
N: 11 / p = 0.028
N: 12 / p = 0.123
N: 13 / p = 0.310
N: 14 / p = 0.043
N: 15 / p = 0.640
N: 16 / p = 0.089
N: 17 / p = 0.095
N: 18 / p = 0.616
N: 19 / p = 0.698
N: 20 / p = 0.531
N: 21 / p = 0.178
N: 22 / p = 0.905
N: 23 / p = 0.224
N: 24 / p = 0.939
N: 25 / p = 0.730
N: 26 / p = 0.184
N: 29 / p = 0.552
N: 30 / p = 0.218
N: 31 / p = 0.079
N: 33 / p = 0.119
N: 34 / p = 0.854
N: 35 / p = 0.671
N: 36 / p = 0.996
N: 37 / p = 0.135
N: 38 / p = 0.351
N: 39 / p = 0.666
N: 40 / p = 0.826
N: 41 / p = 0.277
N: 42 / p = 0.815
Amplitude (uV)
High
Low
Prime Validity
Related Targets
Unrelated Targets

Figure 4.10: *Single-subject voltage (uV) global effect. Each participant includes a participant number (N) and a p-value obtained from the single-subject analysis of variance ANOVA using independent samples. The data for the analyses were obtained by averaging the cluster voltage values of each trial for each condition (high-related, high-unrelated, low-related, low-unrelated) in a time window from 300ms to 800ms as shown in figure 4.6. The interaction in the expected direction corresponds to a greater difference between related and unrelated targets in a high validity condition, relative to a low validity condition.*

Participant	F	p	$\eta_{p}2$	Bf_inclusion
1	2.443	0.120	0.016	0.694
2	0.502	0.480	0.004	0.299
3	0.071	0.791	0.001	0.252
4	3.225	0.074	0.021	0.952
5	0.346	0.557	0.002	0.279
6	0.059	0.808	0.000	0.24
7	0.014	0.906	0.000	0.232
8	1.533	0.218	0.010	0.434
9	0.979	0.324	0.007	0.396
10	2.770	0.098	0.018	0.776
11	4.972	0.028	0.038	2.155
12	2.406	0.123	0.016	0.625
13	1.041	0.310	0.009	0.383
14	4.160	0.043	0.029	1.208
15	0.219	0.640	0.001	0.245
16	2.932	0.089	0.019	0.846
17	2.824	0.095	0.018	0.807
18	0.252	0.616	0.002	0.255
19	0.151	0.698	0.001	0.256
20	0.395	0.531	0.003	0.289
21	1.831	0.178	0.012	0.534
22	0.014	0.905	0.000	0.24
23	1.492	0.224	0.010	0.459
24	0.006	0.939	0.000	0.245
25	0.119	0.730	0.001	0.257
26	1.780	0.184	0.012	0.478
29	0.355	0.552	0.002	0.278
30	1.532	0.218	0.012	0.548
31	3.131	0.079	0.020	0.901
32	0.073	0.787	0.000	0.252
33	2.460	0.119	0.016	0.699
34	0.034	0.854	0.000	0.246
35	0.182	0.671	0.001	0.245
36	0.000	0.996	0.000	0.241
37	2.255	0.135	0.015	0.571
38	0.874	0.351	0.006	0.345
39	0.187	0.666	0.001	0.256
40	0.048	0.826	0.000	0.322
41	1.188	0.277	0.008	0.414
42	0.055	0.815	0.000	0.269

Table 4.3: *Single-subject global effects. The table shows the results of the independent samples ANOVA on each participant's time window from 300-800ms, the data for the analyses were obtained by averaging the cluster voltage values of each trial for each condition (high-related, high-unrelated, low-related, low-unrelated). The table includes the F values, p values, effect size (n_p^2), bayes factor (inclusion) contrasting the alternative against the null hypothesis. None of the participants show a significant interaction in the expected direction, which is greater difference between related and unrelated targets in a high validity condition, relative to a low validity condition.*

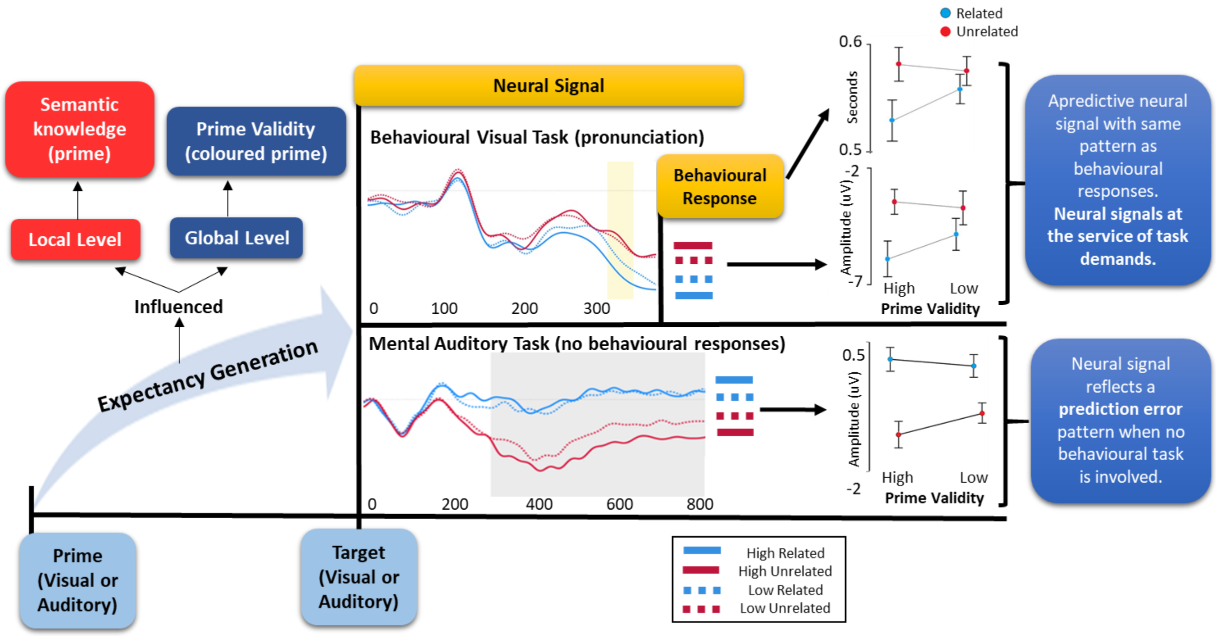


Figure 4.11: *Visual and auditory experiments graphical representation. The model begins at the left with the prime presentation that initiates the generation of an expectation, which is influenced by the individual's semantic knowledge on a local level; and by the prime validity associated with the prime, which is given to the participant as a coloured prime representing either a high or low validity (80% or 20% probability of a related target). Next, the neural signal generated after the target presentation showed two outcomes in the visual and auditory experiments; first, the visual task shows a significant interaction (as shown in figure 4.7) with an apredictive pattern as related targets show more extreme values with respect to related target. The visual task includes behavioural responses (word pronunciation) that show the same pattern as the neural signal, providing evidence for neural signals operating to meet task demands. Second, the auditory experiment included a mental task with no associated behaviour and the neural signal shows the opposite pattern as the visual experiment, which reflects the prediction error elicited by the target, as unrelated targets show more extreme voltage values with respect to related targets. Therefore, the absence of a motor task in this task allows measuring purer cognitive processes, and its presence may influence the neural signal.*

Discussion

The present experiment corresponds to an auditory manipulation of a relatedness proportion paradigm in a semantic priming task, that stems from the visual behavioural experiment investigated by Hutchison (2007). In previous chapters of this thesis, we replicated the author's behavioural semantic effects (priming and relatedness proportion effects) and we then investigated the neural correlates of this visual semantic priming manipulation. These ERP results laid the foundations for the design of the present experiment. The primary aim of this study is to provide an electrophysiological diagnostic tool that can be used to detect neural markers of residual language comprehension in patients that have been diagnosed with a disorder of consciousness, and thus remain unaware of themselves and unresponsive to the environment (Royal College of Physicians, 2020; Zeman, 2001). The lack of responsiveness makes the clinical assessment of patients' cognitive abilities a great medical challenge, as patients cannot perform motor responses (Owen, 2008); hence, the evaluation using electrophysiological measures requires careful task design, where neural markers of conscious processing can be detected and detached from automatic processing, such as it was achieved in the local-global paradigm (Bekinschtein et al., 2009; Faugeras et al., 2012).

In order to detect residual language comprehension in DOC patients, we investigate in a group of healthy participants the neural markers for the generation of strategic semantic expectations while performing this task. According to the target analyses, our pre-registered results show an early 'local' ERP effect around 330ms at parietal electrodes, showing a predictive signal with more extreme values for unpredicted targets (related) relative to predicted targets (unrelated) as it was expected. This effect reflects the violation of local expectations as it is not influenced by the global context (prime validity). On a global level, we detect a late ERP effect in a later time-window (300-800ms) as revealed by our pre-registered follow-up voltage average interaction, which shows an overall predictive signal being more extreme for unexpected targets with respect to expected targets; in addition, the ERP amplitude difference between expected and unexpected targets is greater in a high validity context (80% related targets) where the error is larger for globally unexpected targets, than this difference in a low validity context (20% related targets).

We generate the present auditory experimental manipulation from the results of our previous EEG Visual Experiment, in which we report an early 'local' ERP effect (around 250ms fronto-temporal electrodes) showing a predictive pattern as the ERP amplitudes are greater for unexpected relative to expected words. In this auditory EEG experiment, we also report an early 'local' ERP effect reflecting a predictive ERP pattern consistent with our previous visual local effect. Both results are in line with previous predictive coding accounts where the ERP signals are attributed to the detection of the prediction error response (Dehaene & Christen, 2011; Rohaut et al., 2015). This predictive effect is consistent with N400 effects that have been broadly reported in studies investigating word relatedness by presenting word-pairs either visually or auditorily (Lau et al., 2013; Grossi, 2006; Kutas & Hillyard, 1980). For example, a study by Holcomb and Neville (1990) investigated whether presenting related and unrelated words-pairs visually and auditorily in the same group of participants, would show similar ERP patterns that could account for shared semantic priming mechanisms for both modalities. The results showed greater ERP amplitudes for unrelated targets with respect to related targets in both modalities as expected; moreover, this effect was greater, earlier, and longer in the auditory relative to the visual modality.

The early 'local' effect from our EEG Visual Experiment (Figure 3.2, A and B panels) resembles the MMN elicited in the local-global paradigm (Chennu et al., 2013; Bekinschtein et al., 2009; Faugeras et al., 2012) as both effects show a predictive ERP signal and a scalp topography at fronto-central temporal electrodes, even though both are elicited by clearly distinct stimuli (visual words; auditory tones). In contrast, the early 'local' effect from our EEG auditory Experiment (Figure 4.5) is similar to the classic N400 effect that consistently shows a predictive signal, which is enhanced for unexpected words compared to expected words, and this effect shows a centro-parietal topography (Cruse et al., 2014; Lau et al., 2013; Grossi, 2006; Kutas & Hillyard, 1980). The N400 effect has been attributed as a prediction error signal (Rabovsky & McRae, 2014), where the voltage increases are attributed to the neural cost of processing the error; and a reduced signal for expected items, as a result of semantic facilitation caused by expectation fulfilment (Kuperberg & Jaeger, 2016). Similarly, the MMN has also been proposed as a neural signal that reflects the detection of the prediction error (Chennu et al., 2013; El Karoui et al., 2015). Our early 'local' auditory effect coincide with the N400 effect regarding the EPR pattern and scalp topography distribution; however, although the cluster mass method does not allow for inferences about time for significant clusters, by visual inspection the 'local' effect start to diverge earlier than 200ms, showing an initial mismatch detection that then extends in time showing an N400 pattern. This finding aligns with the theory proposed by Bornkessel-

Schlesewsky and Schlesewsky (2019) in which the MMN and the N400 component are considered as part of the same group of hierarchical effects reflecting underlying predictive processing, and the N400 is viewed as a continuation of the MMN, suggesting that its longer latency reflects neural processing of more complex stimuli than the MMN (Bornkessel-Schlesewsky & Schlesewsky, 2019).

In our previous EEG visual experiment, we also report a later 'Global' ERP effect (Figure 3.2, panels C and D) around 350ms at parieto-occipital electrodes with an apredictive pattern, and as was suggested in the previous chapter, this effect can be either a result of precision-weighted prediction error (Kok et al., 2012) or represent the routing of sensory signals that are oriented to task-direct behaviour due to task demands (e.g., pronunciation task) (Zylberberg et al., 2010). This later effect also yielded a follow-up voltage interaction (Figure 3.3) in which the ERP signal shows a greater difference for related and unrelated targets (priming effects) that were presented under a high validity context, relative to a low validity context (relatedness proportion effects); however, the interaction is not in the expected direction as it showed an apredictive pattern (i.e. extreme voltage for expected targets relative to expected targets), contrary to previous studies that report a predictive ERP relatedness proportion interaction (Lau et al., 2013; Brown et al., 2000). An interesting finding is that the visual 'Global' voltage interaction follows the same pattern as behavioural responses (Figure 3.3), giving rise to the explanation for sensory signals being routed to meet task demands (Zylberberg et al., 2010), which was previously mentioned.

The results of the EEG Auditory Experiment reveal a later 'Global' ERP effect (Figure 4.7) between 300ms and 800ms that involves the global context, as the ERP signal shows a greater difference between related and unrelated targets (semantic priming) presented in a high validity context relative to a low validity context. As our post-hoc analyses show, this 'Global' ERP interaction is not detected prior to 300ms (time window: 100-300ms). The local ERP effect (172-500ms) that only involves priming effects occurs earlier than the global effect (300-800ms) that involves prime validity effects, suggesting that these results could evidence a two stage error profile.

The 'Global' effect ERP pattern is different for our visual and auditory experiments because in the visual global effect (Figure 3.3) the signal is more extreme for the most predictable targets (i.e., high – related); and the signal for the auditory global effect (Figure 4.7) is more extreme for the least predictable targets (i.e., high – unrelated). Regarding the visual global effect, a greater signal for the most predictable targets could be reflecting that individuals actively attend to targets where local or global expectations are met (i.e., Low – Related and High – Related conditions, respectively) in order to meet task demands (i.e., faster pronunciation). Therefore, the signal could be interpreted as upweighting stimuli by precision (Barascud et al., 2016; Heilbron & Chait, 2018), and by placing attentional resources on expected trials individuals would improve the precision of their expectations in these trials (Hohwy, 2012). Although the auditory 'Global' effect (Figure 4.7) shows a similar interaction to the visual 'Global' effect, in the sense that there are greater priming effects in a high validity context relative to a low validity context; the ERP signal for the auditory effect shows larger voltage values for unexpected than expected targets as it was observed in previous language studies (Lau et al., 2013; Brown et al., 2000). Thus, both ERP effects (visual and auditory) show the global interaction, but the signal is flipped with respect to each other, which could be explained by the presence of a motor task. In the presence of a motor task in the visual experiment, the apredictive signal seems to be routed at the service of behaviour as it shows the same pattern as behavioural responses (Zylberberg et al., 2010). Furthermore, in the auditory experiment there is a mental task and no motor task producing behaviour, so the signal reflecting cognitive processing shows a more traditional prediction error response (Dehaene & Christen, 2011; Rohaut et al., 2015), see figure 4.11.

If we consider the ERP signal as reflecting prediction error responses (Dehaene & Christen, 2011; Rohaut et al., 2015), we observe in the auditory 'Global' effect (Figure 4.7) that the signal generated from our four experimental conditions follows an order given by predictability, which accumulates the local and global errors in one effect. For example, in figure 4.6 we can first observe the 'local' error as ERP semantic priming effects (i.e., difference between related and unrelated) where voltage values are greater for unexpected targets relative to expected; furthermore, we can simultaneously observe that the ERP priming effects are greater in the high validity context relative to the low validity context. However, as the global effect is only detected after 300ms and the local effect is detected earlier, these results could be reflecting a two stage error profile, but the absence of an interaction prior to 300ms was detected in exploratory post-hoc analyses, so future research could further investigate whether this tasks reflects a cumulative error or a two stage error effect.

The visual experiment shows distinctive ERP effects for the ‘local’ and ‘global’ contexts with a distinct latency and scalp topographies, as it is observed in the local-global paradigm ERP results. Here, an early ‘local’ error (i.e., MMN response) reflects the use of local context for expectancy generation followed by a later ‘global’ error signal (i.e., P3b response) suggesting the generation of expectations based on the global context; thus, detaching automatic from consciously controlled processing (Bekinschtein et al., 2009; Faugeras et al., 2012). Instead, the auditory experiment seems to show an ERP signal that reflects a cumulative error signal as ‘local’ and ‘global’ effects are embedded within the same effect that starts early and is prolonged in time. As we are proposing this task as a tool to assess residual language processing in DOC patients, by detecting the use of strategic semantic expectations, we would expect to observe this cumulative error as the highest level of a semantic processing hierarchy. Here, individuals use the global context given by the prime validity cue (i.e., type of voice), and the resulting voltage interaction would be a marker of strategic semantic expectations when processing language stimuli, and thus the conscious processing of the task. While only few patients that preserve residual consciousness would be able to follow the task and show the auditory ‘global’ ERP effect; others would not be able to follow the task but may still show signs of language processing. For these patients, we would expect to only observe the auditory ‘local’ ERP effect, where there is a differentiation between the processing of related and unrelated words accounting for some residual language processing; although we cannot claim that this effect involves consciously controlled strategic language processing.

The presence of residual language processing (i.e., presence of N400 effects, auditory 'local' effect in this task) may have prognostic value, as previous studies have shown that their detection in DOC patients could predict their later recovery (Steppacher et al., 2013). In a study by Rohaut and colleagues (2015) N400 effects were detected in 5 patients diagnosed as MCS, where 3 of them regained consciousness later on; moreover, the authors detected an N400 effect in one VS/UWS patient who also recovered consciousness. Another study by Steppacher and colleagues (2013) using a t-CWT algorithm (for more details see Bostanov & Kotchoubey, 2006) detected N400 effects in 32% of VS/UWS patients and 41% of MCS patients; moreover, the authors analysed each patient's ERP by visual inspection detecting N400 effects in 16% of VS/UWS and 21% of MCS patients. According to the authors, patients that evidence N400 effects show a high likelihood of recovering (reported ratio: >100 for MCS, and 22 for VS/UWS). From their methods, visual inspection showed to be more precise in detecting the patients that show N400 effect and later recover (Steppacher et al., 2013).

Regarding the prime ERP analyses (time window 1240 post-prime) that aimed at detecting the generation of expectations from the prime in a high validity context, relative to a low validity context, no significant clusters were identified. Moreover, the cluster analysis failed to detect a difference in the slow wave between prime and target presentation, in both the pre-registered (0-1240ms) and exploratory follow-up (248-1240ms) analyses. In addition, there was no distinctive processing detected in the alpha (8-12Hz) and beta (13-30Hz) time-frequency bands. In our previous visual EEG study, we did not detect any of these expectation markers prior to target onset. One of the explanations provided in the previous chapter (Chapter 3 in discussion) is that we do not have the power to detect these effects, as all of our previous studies were powered to the post-target effect. From two experiments using either visual or auditory language stimuli, we failed to provide evidence that the expectations generated from the primes were different between a high validity context and a low validity context. Nevertheless, the absence of this difference between both conditions, does not necessarily mean that individuals are not generating expectations from the primes. According to de Lange, Heilbron and Kok (2018), one of the misconceptions about predictions is understand them as expecting or predicting something in the future, when they should be understood as a statistical concept, where the prediction stems from a model by inferring potential observations based on this model (de Lange et al., 2018). This process operates under a Bayesian inference conception, where the expectation (i.e., prior) is generated from the model and the context, which is then contrasted with the actual stimuli. Information that is not contained in the contrast between prior and actual stimuli is send up the cortical hierarchy as prediction error (Hohwy, 2017). Hence, probably the most robust neural activity occurs when the prediction is violated, as we can observe in target ERP analyses; in contrast to prime analyses, where there are probably expectations developing from the prime, but as the contrast with the actual stimuli has not occur yet, it becomes more complex to measure this activity in the ERP signal (de Lange et al., 2018; Hohwy, 2017).

Another explanation for the lack of prime effects that we suggested in the previous chapter, referred to the idea that the brain seeks to minimize the use of resources by updating its internal models, so then it would minimise prediction error in future percepts

(Friston, 2010). Therefore, in this paradigm a way of accomplishing this optimization of resources may be predicting related targets, regardless of the prime validity cue (Vidal-Gran, Sokoliuk, Bowman, & Cruse, 2020). As we are dealing with related word-pairs, where a define set of words can be expected; and unrelated word-pairs, where it is not possible to generate expectations as there are a vast amount of possible words that could be unrelated to the prime. Thus, predicting a related target in a low validity context, even though the target has only 22% probabilities of being related, seems more efficient as it is not possible to predict an unrelated target, or at least is challenging to achieve precision in the expectation. Future research could use participant's reports on their expectancy generation process, to investigate how individuals generate expectations from the primes. Moreover, even though we provided two arguments to explain the lack of prime effects in this paradigm, the prime ERP analysis was close to reach significance (smallest cluster $p = 0.055$); therefore, future studies could further investigate these prime effects and perform power analyses to detect them. Furthermore, the post-prime ERP signal shows more negative values for the high validity condition, where expectations are stronger and can reach a higher precision due to top-down influence. This pattern fits with previous studies that have reported the contingent negative variation (CNV), which has been described as an anticipatory slow drift reflected in the ERP signal over fronto-central electrodes, reflecting top-down involvement in the generation of expectations (Chennu et al., 2013; Brown et al., 2008; Walter et al., 1964). Further research would be beneficial to clarify the relationship between the generation of expectations from the prime - involving top-down expectations

- and an eventual CNV before target presentation, that could account as a marker of expectancy mechanisms that rely on top-down processing (Chennu et al., 2013).

Furthermore, participants were requested to guess the upcoming targets, so we measured their confidence ratings in some trials. As a result of using the global context cues and as we expected in our pre-registered hypotheses, individuals were significantly more confident in their guesses for the targets when the primes were presented in a high validity context, relative to their guesses in a low validity context. The use of the global context is also supported by the strategy self-report measure, in which 97.5% (39/40) of participants reported using a strategy while performing the task, which involved using the type of voice cue to accurately predict the target.

As depicted in figure 4.9 and table 4.1, 37 out of 40 participants show a significant difference between related and unrelated targets in an early time-window from 100-500ms at the single-subject level. From these 37 participants, 34 exhibit an N400 ERP pattern, meaning more extreme negative values for unrelated targets relative to related targets (Cruse et al., 2014; Lau et al., 2013; Grossi, 2006; Kutas & Hillyard, 1980). The local effect -N400 ERP effect- is detected here at both group and single-subject analyses, being therefore suitable to be detected in individual DOC patients as other studies have proposed (Steppacher et al., 2013; Rohaut et al., 2015; Cruse et al., 2014). However, the global effect -cluster voltage interaction- was only successfully detected at the group level in a time-window from 300-800ms as shown in figure 4.7, and none of the participants show evidence for a voltage interaction at the same time-window in the single-subject analyses, see figure 4.10 and table 4.2. The lack of a global effect on a single-subject level, questions this effect's clinical utility for detecting strategic semantic processing in DOC patients, however, these results may be due to specific analyses parameters, such as the time-window selection or channels of interest. Therefore, future research should be directed into investigating this global effect and exploring other methods/analyses that could either detect or fail to provide evidence for this effect at a single-subject level. In the following section, we will present two patients (i.e., patient cases) that were assessed with this auditory relatedness proportion paradigm here proposed. The analyses that are presented in this chapter aim to investigate the ERP responses that are elicited by this task across a population of healthy individuals. The task responses are intended to reflect hierarchical language processing, where the absence of ERPs can provide information about the patient's perception abilities (Beukema et al., 2016); next, finding the auditory 'local' effect would provide evidence for some semantic processing. At the highest complexity level, the presence of a later 'Global' interaction (global effect) would represent strong evidence that the patient is strategically following the task. Strategic involvement in this task is based on healthy participants' previous results, including the behavioural interaction previously observed (RPEs), with its supporting neural markers previously mentioned, besides the support from self-report measures indicating the use of strategy while performing the task.

The patient's data, which is analysed on a single-subject basis is then compared to the group results as it is common practice (e.g., Bekinschtein et al., 2009; Cruse et al., 2012; Beukema et al., 2016). It must be considered that ERP single-subject analyses are more likely to show a rough signal as opposed to a smoother signal obtained in group analyses, where the smoother signal is a result of the averaging process across participants, as averaging reduces the noise in the signal and reveals a closest estimate of the signal elicited by a certain stimulus (Luck, 2014); therefore, single-subject results should be treated with cautious.

EEG Patient Cases

Methods

Participants

We recruited 3 patients from two specialist neurorehabilitation centres in central England, however, one patient data was excluded due to excessive muscle artefact. Therefore, we report two patient's data. Patient 1 was 21 years old at the time of testing and 20 years old at the time of injury. The patient was diagnosed with UWS/VS which was a result of hypoxic injury from cardiac arrest. Patient 2 was 56 years old at both the time of injury and testing and received the MCS+ diagnosis due to unknown cause. Diagnoses were taken from each patient's clinical SMART assessments (Gill-Thwaites & Munday, 2004). The study was approved by the West Midlands Coventry and Warwickshire Research Ethics Committee and was sponsored by the University of Birmingham, England. Consultees of each patient were identified by the clinical team and approached to provide written consent.

Stimuli

We used the same stimuli as the auditory EEG experiment.

Procedure

Stimuli were displayed to participants using Psychtoolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) in Matlab (Mathworks, Inc., Natick, Massachusetts). We use a similar procedure as with the group of healthy participants. Each patient was tested individually, the experiment includes 352 trials and we only presented the ERP trials described in EEG experiment 2, as patients are not able to provide behavioural responses. Patients only heard the words through the headphones, and we did not present visual stimuli on a screen. Each trial started with silence for 750ms; next the prime word was played through the headphones in either a male or a female voice. Subsequently, 1240ms (SOA) after the prime onset, the target word was played in the same type of voice as the prime. Next, silence remained for 1000ms until the next trial started. We provided simplified auditory instructions to patients as follows: "You are about to hear a woman and a man saying two words at a time. When you hear them speak the first word try to guess what they will say next. The man (or woman depending on assignment) will mostly say words that are related to each other, like hide and seek".

EEG recording

We continuously recorded the EEG signal of both patients at the bedside, with a high-density 128 channel AntNeuro EEG system (AntNeuro b.v., Enschede, Netherlands). The brain activity was acquired at a sampling rate of 500Hz, and we kept impedances below 20 kΩ when possible. We placed the ground electrode on the skin at the left side of the forehead, and we previously sanitised the skin at this area and applied conductive gel. The voltage that was measured by the 127 electrodes on the shielded waveguard cap was referenced to the CPz channel. Moreover, we recorded ECG activity using external sensors. We placed two ECG electrodes, one on the left side of the chest and the other electrode on the right side.

EEG Pre-Processing Pipeline

We low pass filtered the continuous EEG data at 40Hz using the finite impulse response filter implemented in EEGLAB (Delorme & Makeig, 2004). As in our previous analyses we did not perform high-pass filtering due to our interest in analysing slow-waves, we maintain this as we will compare patients and healthy individuals' data. We then segmented the filtered EEG signals into epochs from 750ms before the onset of the prime up to 1240ms post-target. Subsequent artefact rejection was done manually with visual inspection using EEGLAB (Delorme & Makeig, 2004), where we first rejected bad channels and bad trials. Next, to remove stationary artefacts, such as blinks and eye-movements, the pruned EEG data was subjected to an independent component analysis with the runica function of EEGLAB, which artefact components containing eye-blinks, eye-movements and data discontinuities were later rejected, so their contribution was subtracted from the data. Next, we interpolated the data of any previously removed channels via the spherical interpolation method of EEGLAB and re-referenced the data to the average of the whole head.

Analyses

Single subjects' analyses: PDOC Patients

Only the healthy participants' data is included in the pre-registration project, therefore, the patient's data analysis is classified as exploratory. We analyse single subject ERP time-courses for the patient cases, considering the same time-window of interest as healthy participants for the early local-effect from 100-500ms, where we test the main effect of relatedness (Related-Unrelated). The time-windows are compared with the cluster mass method of the open-source Matlab toolbox FieldTrip (Oostenveld et al., 2011), as described in EEG Experiment 1. However, as this analysis is performed on single subjects, we use the participant's individual trials for the ERP analysis, instead of averages across participants as are used for the group-level analyses.

If a patient shows a significant cluster in the local-effect ERP analysis (main effect of relatedness), we will then perform a single subject analysis of variance (similar to EEG experiment 2) to test for the interaction ((High-Related – High Unrelated) – (Low-Related – Low-Unrelated)). For this analysis, we select the data from the channels that compose the significant cluster in the 100-500ms time-window. Next, each trial is averaged across all cluster channels and time-points within the time-window, resulting in one averaged value per trial. The single values will be grouped by condition to run the voltage interaction using Jasp 0.9.1.0 software (JASP Team, 2019).

Exploratory single-subject results

Patient 1

We tested for a main effect of relatedness between related and unrelated targets in an early time window from 100 to 500ms, as indicated by the healthy participants' group analyses. According to the single-subject cluster-based permutation analysis, there was no significant main effect of relatedness of the target (one positive cluster p = 0.755), see figure 4.9. As we found no effect of relatedness, we did not conduct an interaction analysis as specified in the analyses section above.

Patient 2

In the same way as patient 1, we tested for a main effect of relatedness between related and unrelated targets in an early time window from 100 to 500ms. The single-subject cluster-based permutation analysis yielded a significant temporal electrodes main effect of relatedness of the target at approximately 350ms post-stimulus (positive cluster: 260 – 444ms, p = 0.006), which is shown in figure 4.10. Voltage values in this significant cluster are more extreme for unrelated targets, relative to related targets; hence, the effect suggest a predictive signal as it was expected from our pre-registered analyses and also observed in our group-analyses from healthy participants.

The patient showed a significant main effect of relatedness, thus, as it was stated in the analysis section, we test for the voltage interaction between the validity of the prime and the relatedness of the target in the 100-500ms time-window. The results yielded no significant voltage interaction between the validity of the prime and the relatedness of the target (F (1, 30) = 1.133, $p = 0.296$, $\eta p^2 = 0.036$), see Figure 4.11. Moreover, the Bayesian Repeated Measures ANOVA showed anecdotal evidence in favour of the null hypothesis ($BF_{inclusion} = 0.398$).

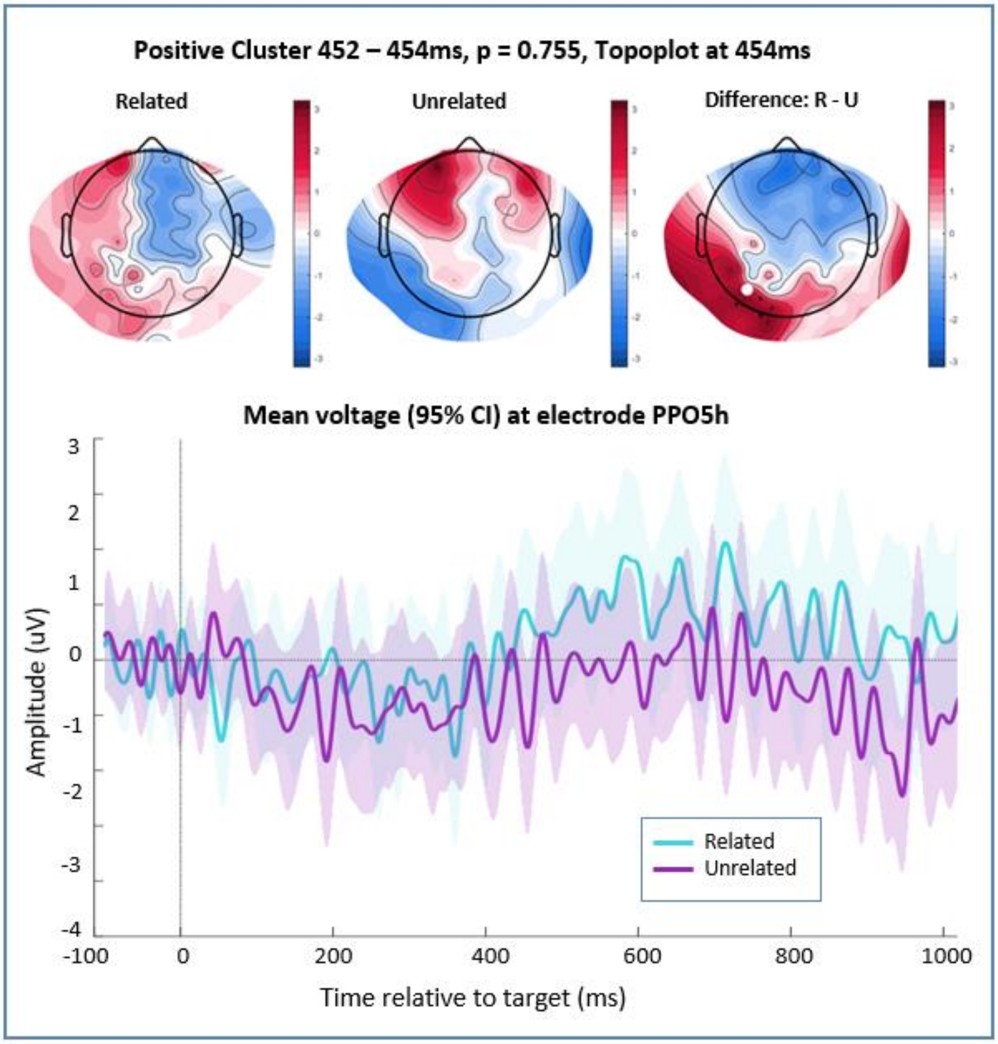


Figure 4.12: *Patient 1 ERP analysis of main effect of relatedness, containing Mean voltage (95% CI) at electrode PPO5h, which analysed the difference between related and unrelated targets in a time window from 100-500ms post target. The analyses failed to detect a significant difference between conditions (positive cluster p = 0.755).*

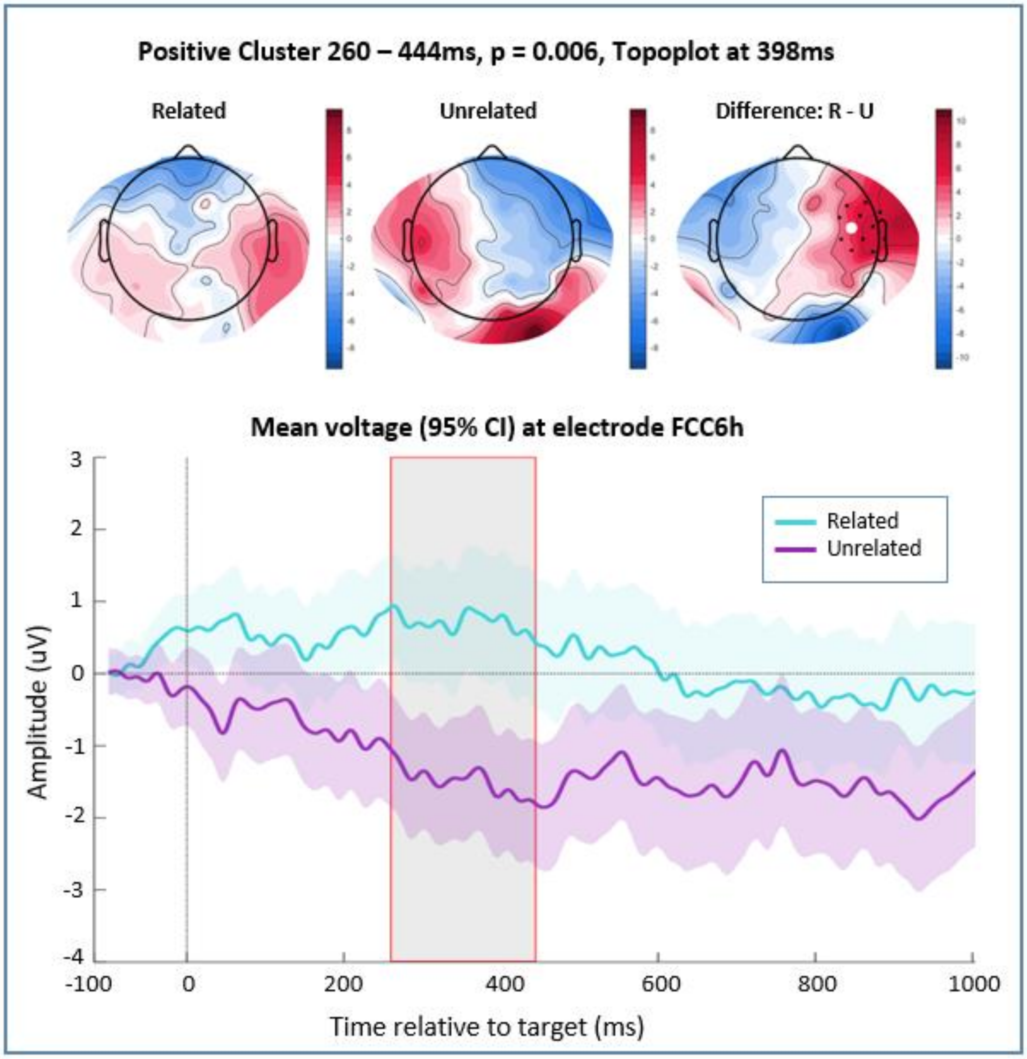


Figure 4.13: *Patient 2 ERP analysis of main effect of relatedness, containing Mean voltage (95% CI) at electrode FCC6h, which analysed the difference between related and unrelated targets in a time window from 100-500ms post target. The analyses failed to detect a significant difference between conditions (positive cluster p = 0.755).*

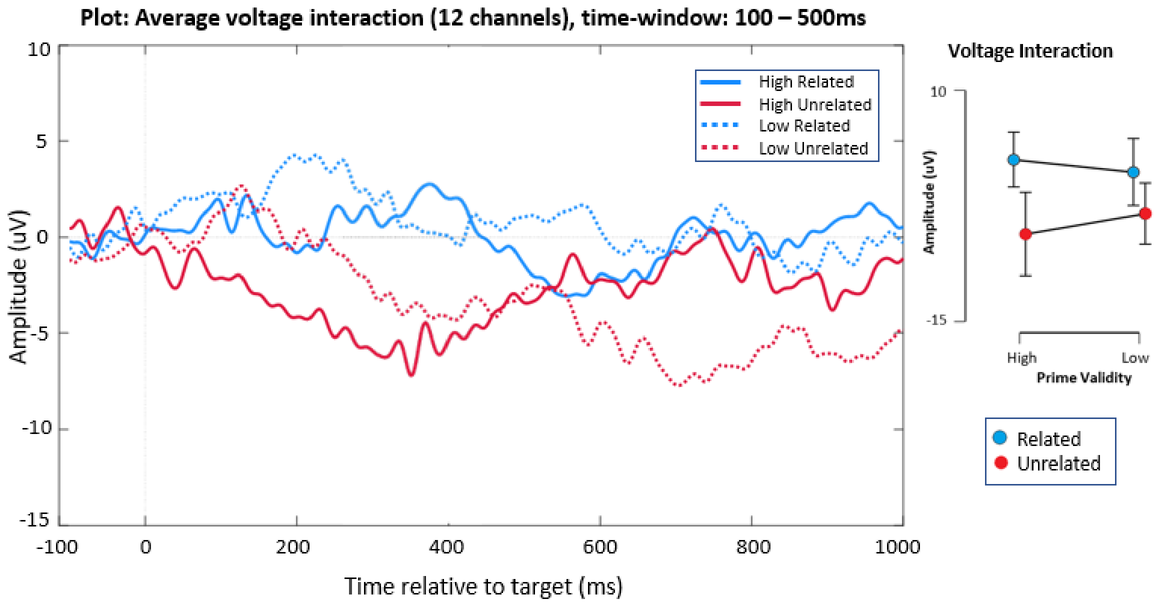


Figure 4.14: *Patient 2 ERP interaction (prime validity * Relatedness of the target), showing the mean of electrodes (12 electrodes) within the cluster from 100 – 500ms., The analyses failed to detect a significant interaction between conditions (positive cluster p = 0.296).*

Discussion

We tested an auditory relatedness proportion paradigm in two patients that were diagnosed with disorders of consciousness, one patient received an MCS diagnosis, and the other patient was classified as UWS/VS. This task aims at detecting evidence for the use of strategic semantic expectations when processing language stimuli, so these expectations can account as a marker of conscious processing. To fulfil this purpose, we provide a within-trial 'local' context given by related or unrelated word-pairs where individuals can generate expectations about the target based on the prime. In addition, we provide a 'global' context across the task that cue participants about the probability of the target being related to the prime (i.e., low and high validity contexts), the cue is delivered to participants in the type of voice that pronounces the primes (i.e., male and female voices) where each type of voice corresponds to a certain probability (e.g. male voice 80% probability of related target; female voice 20% probability of a related target).

As observed in the group of healthy participants, we first expect to find a local ERP effect around 330ms (time window from 100ms to 500ms) with the ERP signal showing a predictive pattern (i.e., enhanced signal for unrelated targets relative to related targets) and a centro-parietal topography, such as previous studies that have reported N400 effects when contrasting related and unrelated targets (Cruse et al., 2014; Lau et al., 2013; Grossi, 2006; Kutas & Hillyard, 1980). If participants display the 'local' ERP effect, we search for the presence of the interaction (prime validity * target relatedness) in the same time window where the local effect is detected. As we noted from healthy participants, the local and global effects represent a cumulative error, therefore we would expect to observe the interaction in a similar time window as the main effect.

In the single-subject patient cases, patient 1 who had a UWS/VS diagnosis, shows absence of local ERP effect, so there was no evidence for a difference in the processing of expected targets relative to unexpected targets. Regarding patient 2, who has an MCS+ diagnosis, the results showed a significant early effect (around 350ms), similar as healthy participants, indicating evidence for a differentiated processing between expected and unexpected targets on a local level. Due to the presence of this effect, we test for the global effect interaction, which did not reached significance. However, if we visually inspect the interaction 'global' effect, it has the same pattern as the 'global' effect in healthy participants showing greater priming effects in a high validity context relative to a low validity context.

In general, is expected that MCS patients show greater responsiveness to language stimuli than UWS/VS patients (Royal College of Physicians, 2020). In Rohaut and Colleagues (2015) 6 out of 29 DOC patients showed N400 effects, where 5 corresponded to MCS diagnosis and only one was diagnosed as VS/UWS. Another study by Beukema and colleagues (2016) only reported N400 effects in one patient that was classified as MCS, while these effects were not detected in any of the patients with a VS/UWS diagnosis. As we can observe in our two patient cases, patient 1 (UWS/VS) shows no indication of target differentiation; in contrast with patient 2 (MCS), who despite not showing the strategic global component of the task, does show evidence for violation of local semantic expectations (Fitz & Chang, 2019; Rabovsky & McRae, 2014); so, we can infer that the patient show a differentiated processing for related and unrelated targets. Given all the evidence from previous experiments and this two patient cases, this auditory paradigm shows proof-of-concept that it can be accounted as a viable tool for measuring hierarchical language processing in patients diagnosed with a disorder of consciousness. Moreover, this task includes the identification of strategic language processing where its electrophysiological markers (i.e., the 'Global' ERP effect) are only observed in healthy participants that consciously follow the global context and report using a strategy for this purpose. However, as the global effect was not detected in healthy participants on a single-subject level, this effect does not currently have clinical utility until further research could clarify whether there is no evidence for the presence of this effect, or if the analyses that we conducted were not sufficient to detect the effect in single-subjects. If the single-subject effect gets detected using other methods/analyses, future studies should aim at replicating the present experiment and test it on a significant patient sample, including several diagnostic labels, such as UWS/VS and MCS. Moreover, studies investigating prognosis in DOC could also benefit from this paradigm, in which the presence of neural markers of strategic semantic expectations could contribute to estimate the likelihood of recovery.

CHAPTER 5: GENERAL DISCUSSION

Although disorders of consciousness (DOC) only affect a small percentage of the population, they have a dramatic impact in patients' and their relatives' lives, as it corresponds to a severe disability where patients lose their autonomy and their interaction with the environment (Royal College of Physicians, 2020). These disorders are a stressful and traumatic experience for the patients' relatives (Fins, 2013), where there is diagnostic and prognostic uncertainty given by the variation between cases, which makes it difficult for clinicians to establish recovery patterns among patients, besides the lack of precise diagnostic tools to measure the patient's level of consciousness (Schnakers et al., 2009; Monti et al., 2010). Many efforts have been devoted to reach a scientific consensus on a definition of consciousness, however, there is still no unified definition across disciplines as it corresponds to a complex human phenomenon (Owen, 2008). The lack of an integrated definition creates difficulties when researchers and clinicians aim to measure manifestations of consciousness in both experimental contexts and clinical settings. For DOC patients, who are behaviourally unresponsive, it is assumed that they are deprived from conscious experiences; although, studies have consistently shown that some patients are able to modulate their brain activity to commands when they are assessed with paradigms using neuroimaging (Monti et al., 2010; Owen et al., 2006) or electrophysiological methods (Faugeras et al., 2012; Cruse et al. 2011; Bekinschtein, 2009).

Despite all the available evidence on covert awareness in DOC patients, that stems from neuroimaging and electrophysiological research, the behavioural examination continues to be the clinical standard procedure in the UK to determine the patient's level of awareness; as it is indicated in the UK clinical guidelines for the management of patients diagnosed with prolonged disorders of consciousness, which states: "electrophysiological tests and more sophisticated imaging techniques (such as fMRI, PET scans etc) do not form part of routine clinical evaluation for patients with PDOC" (Royal College of Physicians, 2020, p. 46). Behavioural scales to assess the level of awareness in patients diagnosed with DOC (e.g. CRS-R) aim at differentiating intentional from reflexive behaviour (Rohaut, Eliseyev & Claassen, 2019). Although these scales are necessary to measure patients' behavioural responsiveness to the environment, they have shown a lack of diagnostic precision due to patients' inability to generate overt behaviour (Schnakers et al., 2009; Andrews et al., 1996; Childs et al., 1993); hence, the importance of demonstrating evidence of covert awareness with neuroimaging or electrophysiological measures in this group of patients (Rohaut et al., 2019). The difficulty lies in differentiating between conscious and unconscious processes; identifying which cognitive processes can account for a full conscious experience; and how we can measure these processes (Seth et al., 2008). Therefore, electrophysiological or neuroimaging studies investigating covert awareness in DOC patients should compare their results with healthy individuals (Rohaut et al., 2019), by measuring specific cognitive processes and evidence a distinctive use of strategic controlled processes rather than non-strategic or automatic brain processes (Faugeras et al, 2012).

Some electrophysiological measures have shown potential diagnostic value, by providing evidence of strategic controlled processing in DOC patients when they are presented with differentiated auditory stimuli, such as the P3b (P300) response to global violations of auditory regularities (Faugeras, 2012; Bekinschtein, 2009); however, a meta-analysis suggested that the P300 lacks prognostic properties with respect to other electrophysiological measures (e.g. oscillatory EEG responses) (Kotchoubey & Pavlov, 2018). ERP Measures, such as the P300 are largely used in DOC research, they have not been successfully introduced into the clinical practice as expected (Royal College of Physicians, 2020). To date, there is no task that has demonstrated evidence for residual language processing in DOC patients involving the use of strategic semantic expectations. The view of expectations influencing language processing has been broadly studied and this view has been supported from predictive coding accounts (Lewis & Bastiaansen, 2015; Ylinen et al., 2016; Lau et al., 2013; Hutchison, 2007; Kuperberg & Jaeger, 2016). The main postulate is that language mechanisms would operate by generating expectations about upcoming semantic stimuli in order to facilitate comprehension (Kuperberg & Jaeger, 2016). Using contextual information and internal models, individuals generate expectations about upcoming stimuli, which are constantly being contrasted with the encountered stimuli; when predictions are met, the new information feeds the model to maintain prediction accuracy for future percepts; while predictions that are violated force the model to adapt to new upcoming information in order to better predict future percepts (Friston, 2010). Neural processing follows a hierarchical organization where there is an interplay between perceptual bottom-up signals at the lowest levels and top-down predictions that are generated from the highest level of the hierarchy; top-down predictions intend to explain lower bottom-up signals and thus reduce prediction error signals generated at these lower levels, to better predict upcoming stimuli (Clark, 2013; Friston, 2005).

The end goal of this thesis is to show proof-of-concept of a paradigm that could be implemented in the clinical practice as a viable tool to assess residual language comprehension in DOC patients. Specifically, this task intends to detect the use of conscious strategic semantic expectations when processing language stimuli, besides accounting for the patient's level of awareness, as language constitutes a fundamental ability when determining that someone is aware of themselves and the environment, meaning that evidence for language comprehension is in itself evidence of consciousness (Owen & Coleman, 2008).

To accomplish this goal, we here present a series of experiments using a language paradigm that provides behavioural evidence for the generation of strategic semantic expectations with its respective neural markers, and then we adapted this paradigm to measure and meet patients' abilities.

Behavioural evidence for strategic semantic expectations

In our behavioural experiments, we conducted a partial replication of a relatedness proportion paradigm in a semantic priming task that was implemented by Hutchison (2007), where the relatedness of the target (i.e., related – unrelated) represents the local context, and trials are presented within a context that represents the global context across the task (i.e., high prime validity – low prime validity). In Chapter 2, we report two behavioural experiments that provide evidence for the use of strategic semantic expectations in this paradigm. On a local level, we observe that individuals are significantly faster to pronounce targets that are related to the prime, in contrast to targets that are unrelated to the prime (i.e., priming effect). This results are consistent with what we expected and with previous findings that have reliably report semantic priming effects when individuals are presented with related and unrelated prime-target word-pairs (Hutchison, 2007; Hill et al., 2002; Perea & Rosa, 2002; Perea & Gotor, 1997; Keefe & Neely, 1990; Neely & Keefe, 1989; Lupker, 1984). On a global level, we provide evidence for an interaction between the validity of the prime and the relatedness of the target, which is driven by related targets being significantly faster when they are presented in a high validity context, relative to a low validity context. This interaction means that individuals show greater priming effects in a context with high prime validity relative to a context with low prime validity (Neely et al., 1989; de Groot, 1984); and this effect indicates that individuals are using the prime cue to engage in consciously controlled expectations to predict the target, by showing faster responses in trials where the contextual cue (i.e., coloured prime) benefits expectation fulfilment (i.e., related targets presented in a high validity context) (Hutchison, 2007; Keefe and Neely, 1990).

Our behavioural results replicate Hutchison's (2007) findings, which provide evidence for RPEs for the long SOA manipulation (1240ms) and failed to provide evidence for this effect at the short SOA (267ms) (Hutchison, 2007). Ours and Hutchison's (2007) findings are consistent with previous studies that have reported RPEs at long SOA manipulations (Grossi, 2006; Stolz et al., 2005; Brown et al., 2000; Neely et al., 1989; Keefe & Neely, 1990; de Groot, 1984); as these effects are present when individuals have sufficient time between prime and target to generate controlled expectations (Hutchison, 2007). The creation of both high and low validity contexts allows for the generation of strategic top-down expectations based on contextual information (Hutchison, 2007; Lau et al., 2013). If individuals would not be generating top-down expectations that rely on the context, we would observe similar priming effects, regardless of the validity context (i.e., high and low) in which they are presented (Hutchison, 2007).

A novel finding from this thesis is that the RPE (i.e. prime validity interaction) in Behavioural Experiment 2 is significantly modulated by the use of strategy, as participants that report using the prime strategically (i.e. Strategy group) show a significant RPE, whereas individuals that report not using the prime strategically (i.e. No Strategy group) failed to show an RPE and only exhibit significant priming effects (i.e. faster responses for related than unrelated targets) that are not modulated by the prime validity. The modulation of the RPE by the use of strategy, provides direct evidence that individuals require to generate effortful expectations to show RPEs (Vidal-Gran, Sokoliuk, Bowman, & Cruse, 2020). In order to use the global context strategically, individuals require to direct their attentional resources to the task and actively engage in the generation of expectations; in turn, carrying out these actions requires the presence of full conscious processing (Seth et al., 2008), showing that this task serves the purpose of detecting conscious language processing. In addition, these results are consistent with Hutchison's (2007) attentional control (AC) measures, as these revealed to modulate the size of the RPEs, meaning that individuals with higher AC scores showed greater RPEs; this findings support the involvement of strategic processes on RPEs, as attentional resources should be directed to the task in order to show these effects (e.g., low attentional control individuals did not show RPEs) (Hutchison, 2007).

Neural markers of strategic semantic expectations

Once we provided behavioural evidence for the involvement of strategic expectations in a relatedness proportion manipulation by Hutchison (2007), we conducted the same experimental manipulation as Behavioural Experiment 2, and we included electrophysiological recordings to measure the neural markers of strategic semantic expectations. First, we replicate the behavioural RPE as previously reported in Behavioural Experiments 1 and 2. Moreover, we detect an early 'local' ERP effect at fronto-temporal electrodes around 250ms that shows a predictive pattern, as the signal shows more extreme values for unexpected targets (i.e., unrelated) relative to expected targets (i.e., related), as we expected in our hypotheses. The 'local' ERP effect is followed by a later 'global' ERP effect at parieto-occipital electrodes around 350ms that shows an apredictive pattern, where values are more extreme for expected targets (i.e., related) relative to unexpected targets (i.e., unrelated), contrary to what we expected. This effect shows an interaction with the global context, where there is a greater ERP signal difference between related and unrelated targets presented in a high validity context, relative to this difference in a low validity context, although the signal shows an apredictive pattern (i.e., more extreme values for expected targets).

Overall, the results of this visual EEG experiment provide evidence for a two-stage expectation profile, which is similar to the two-stage profile reported in the local-global paradigm (Bekinschtein et al., 2009), where an initial MMN is detected when expectations are violated on a local level, followed by a later positivity (P3b) reflecting the violation of expectations that are built from the conscious use of contextual information (Bekinschtein et al., 2009; Faugeras et al., 2012). In our results, the early 'local' ERP effect reflects a prediction error signal showing a greater ERP signal for targets where local expectations are violated (i.e. unrelated targets) relative to targets where local expectations are met (i.e. related targets); similar to the predictive signal and fronto-central topography observed in the MMN, although our effect shows a longer latency than the MMN (Chennu et al., 2013; Bekinschtein et al., 2009; Faugeras et al., 2012; El Karoui et al., 2015). Moreover, the local effect was source localised to the right middle frontal gyrus, which is similar to what Rohaut and colleagues (2015) reported, where they localised the N400 effect to the left middle frontal gyrus; which has been previously linked with semantic categorisation (Noesselt et al., 2003), although as we mentioned in Chapter 3, this region was not included in our pre-registered analyses, so future studies could investigate the middle frontal gyrus involvement in this task or similar experimental manipulations.

Next, we detect a late 'Global' ERP effect that interacts with the global context (i.e. prime validity manipulation), so as the P3b in the context of the local-global paradigm; however, our effect shows an apredictive pattern, contrary to the P3b that shows a predictive signal (Bekinschtein et al., 2009). The apredictive signal (i.e., more extreme for expected targets) that we observed in our interaction is inverted with respect to the ERP interaction found in previous relatedness proportion manipulations, which have reported a modulation of the N400 effects given by the global context (Lau et al., 2013; Brown et al., 2000). For example, the study by Lau, Holcomb and Kuperber (2013) implemented a relatedness proportion paradigm in a semantic priming task, where they presented prime-target word-pairs in two experimental blocks; one block had a high proportion of related word-pairs (50%) and another block containing a low proportion of related word-pairs (10%). The results showed a greater N400 effect (i.e., predictive signal) to targets presented in a high proportion, relative to targets presented in a low proportion (Lau et al., 2013).

One explanation for our apredictive signal is that it may be produced by individuals placing their attentional resources on trials where both local and global expectations are fulfilled (e.g. related trials on a local level and related trials under a high validity context on a global level), which would generate a larger signal for these trials (Barascud et al., 2016; Hohwy, 2012), with respect to trials were expectations are locally and globally violated; similar to previous studies that have reported a larger signal for predictable stimuli in contrast to unpredictable stimuli (Barascud et al., 2016). This attentional explanation is supported by up-weighting expected stimuli by precision (Heilbron & Chait, 2018), where attending to expected stimuli increases the precision in the expectations, which in turn can be reflected in an increased ERP signal (Hohwy, 2012). However, it should be noted that previous studies that have reported ERP relatedness proportion effects (Lau et al., 2013; Brown et al., 2000) or even the global effect in the local-global paradigm (Bekinschtein et al., 2009; Faugueras et al., 2012) have not included pronunciation responses in each trial as we are presenting here. Therefore, our apredictive signal can also be explained as neural signals being routed to meet task demands, as the ERP signal follows the same pattern as motor behaviour (Zylberberg et al., 2010), which can be seen in Figure 3.3, where both the behavioural RPE and ERP voltage follow the exact same pattern. Support for this explanation stems from our source localisation analyses, that localised the 'Global' ERP effect to the supplementary motor area, which has been associated with speech motor control, verbal working memory and semantic top-down involvement (Hertrich et al., 2016); thus, reinforcing the idea of the global effect as sensory signals routing behavioural responses. However, as we mentioned in the local effect above, this source localization finding requires future investigation as this region was not part of our regions of interest when we defined our experimental hypotheses.

Even though our global effect shows an apredictive signal, we detected an interaction that involves the use of the global context, suggesting that individuals were using the prime strategically to successfully predict upcoming targets (Hutchison, 2007); which was also confirmed by individuals' self-report, where all participants that were included in the analyses, reported using the colour of the prime strategically while performing the task. Therefore, our late 'Global' interaction suggest the involvement of strategic expectations, in a similar way as the P3b responses are only present when individuals are aware of the global context (Bekinschtein et al., 2009; Faugeras et al., 2012).

Strategic semantic expectations in the context of DOC

We designed an auditory EEG experiment that is based on our previous behavioural and neural evidence for the use of strategic semantic expectations in a relatedness proportion paradigm (Hutchison, 2007). In order to measure markers of strategic expectations in DOC patients, we adjusted the experimental design to meet patient's abilities by presenting words auditorily, where the prime validity cue was given by the type of voice (i.e., either female or male voice); we instructed participants to do a mental task by guessing the upcoming target when they hear the prime; and we provided information about the global rule (i.e., prime validity manipulation). The results from healthy participants yielded an early 'local' ERP effect at centro-parietal electrodes around 350ms that reveal a predictive signal, as locally unexpected targets (i.e., unrelated targets) show more extreme voltage values, relative to locally expected targets (i.e. related targets). This effect is consistent with previous studies that have reported N400 effects in semantic priming tasks, showing greater negativity for unrelated targets relative to related targets, and this effect shows a centro-parietal topography (Cruse et al., 2014; Lau et al., 2013; Grossi, 2006; Kutas & Hillyard, 1980). From a predictive coding perspective, the N400 effect has been proposed as a semantic prediction error signal, that is observed when semantic expectations are violated relative to when these expectations are met (Rabovsky & McRae, 2014). Here, the ERP signal shows greater voltage values for unexpected targets in contrast to expected targets (Fitz & Chang, 2019; Rabovsky & McRae, 2014).

Furthermore, we report a late 'Global' ERP effect at parietal electrodes between 300ms and 800ms that shows a significant interaction with the global context (i.e., prime validity), evidencing a greater voltage difference between related and unrelated targets that were presented under a high validity context, relative to this difference when presented in a low validity context. Moreover, the signal shows a predictive pattern as the voltage values are more extreme for unexpected targets (i.e., unrelated targets), with respect to expected targets (i.e., related targets). The results are consistent with previous studies that have reported a modulation of the N400 effects driven by the relatedness proportion manipulation (i.e., high prime validity in our experiment), where the high proportion condition shows larger voltage differences between related and unrelated targets, relative to a low proportion; and the signal shows a predictive pattern (Lau et al., 2013; Brown et al., 2000). The greater difference observed in the ERP signal between related and unrelated targets (i.e., priming effects) in a high validity context (relative to a low validity context), is a result of individuals engaging in conscious expectations about the target that depend on contextual information (e.g., prime validity) (Lau et al., 2013; Hutchison, 2007). When these top-down expectations are violated (e.g., unrelated target in a high validity) the prediction error is greater in a high validity context, in which individuals expect to encounter a related target; and this error involves a neural cost that is reflected in an ERP signal increase with respect to targets where the expectations are met (e.g., related targets in a high validity context) (Lau et al., 2013).

The results from healthy individuals in this auditory version of this relatedness proportion paradigm (Hutchison, 2007) show a two-stage expectation profile, as we observe an early effect that only involves the local context (i.e., related and unrelated targets), followed by a late effect that interacts with the global context (i.e., prime validity * target relatedness). The local effect is equivalent to the classical N400 effect that is largely reported in language studies (Cruse et al., 2014; Lau et al., 2013; Grossi, 2006; Kutas & Hillyard, 1980) and the global effect is equivalent to the N400 modulation by contextual information that has been previously reported (Lau et al., 2013; Brown et al., 2000). However, our early effect start to diverge before 200ms by visual inspection, indicating an initial mismatch detection that continues towards an N400 pattern. This effect is consistent with Bornkessel-Schlesewsky and Schlesewsky (2019) theory, where the MMN and N400 are part of the same group of effects that reflect predictive processing, but their latency would index stimulus complexity (Bornkessel-Schlesewsky & Schlesewsky, 2019). Therefore, we would expect to observe as a 'local' effect in DOC patients the initial mismatch followed by the N400 effect. Following this early effect, we would expect to find an interaction that involves the global context, therefore, suggesting the use of strategic expectations (Hutchison, 2007; Lau et al., 2013).

Furthermore, our results show that participants were significantly more confident in their guesses about the target, when the primes were presented in a high validity context relative to a low validity context; and all participants, except from one, reported using the primes strategically throughout the task. Both findings provide evidence that individuals were using the primes strategically and generating expectations based on the global context in order to better predict upcoming targets (Hutchison, 2007).

The global effect of the auditory manipulation shows a predictive signal and does not have a pronunciation task in each trial as our previous visual study, which showed an apredictive signal in the global effect; this difference provides evidence for the explanation given above about neural signals routing behavioural responses when individuals have to meet task demands (Zylberberg et al., 2010). The observed difference in the direction of the ERP signal between both studies is a novel finding from this thesis, as it raises the question whether including a motor task (e.g. button press, pronunciation, etc.) when investigating prediction mechanisms could confound ERP results; as the ERP signal could be influenced by the processing of task demands (Zylberberg et al., 2010), instead of reflecting purer cognitive processing, as in this case a prediction error response (Dehaene & Christen, 2011; Rohaut et al., 2015), when only including a mental task in the experimental design. Future research could focus on investigating the direct influence that task demands may have on neural signals, as many tasks aiming to measure cognitive processing include behavioural responses (e.g., button press), which could influence the neural signals and confound results. These results challenge a prediction/prediction-error interpretation for all evoked potentials (Friston 2012; Friston & Kiebel, 2009), where later processes might use the earlier information provided by predictive mechanisms (e.g., early effect in visual experiment) to produce behaviour (e.g., late effect in visual experiment) or represent the information for consciousness.

Clinical implications of strategic semantic expectations in DOC patients

We tested the auditory task in two patients with prolonged disorders of consciousness; the first patient had an VS/UWS diagnosis, whereas the second patient had an MCS+ diagnosis, both patient's data were analysed with single-subject ERP analysis. Patient 1 failed to show a 'local' ERP effect, which means that there was no evidence for a differentiated processing of related and unrelated targets (Cruse et al., 2014; Grossi, 2006; Kutas & Hillyard, 1980). The data analysis of Patient 2 yielded a significant 'local' ERP effect at temporal electrodes around 350ms showing a predictive ERP signal, where unexpected targets (i.e., unrelated) have greater voltage values relative to expected targets (i.e. related). This effect implies that the patient can process the meaning of speech at some level that probably does not involves conscious processing but may provide some directive for a residual neural architecture for semantic processing, that would support the emergence of language abilities in an eventual recovery (Beukema et al., 2016; Rohaut et al., 2015; Cruse et al., 2014). Future studies should further look at what this specific effect means for recovery, as several previous studies suggest that is more likely to detect N400 effects in MCS patients with respect to VS/UWS patients (Beukema et al., 2016; Rohaut et al., 2015), which is also consistent with our results, as the local effect was observed in Patient 2 that has a MCS diagnosis, and not in Patient 1 that has a VS/UWS diagnosis. Moreover, MCS patients have a higher level of responsiveness relative to VS/UWS, and therefore are more likely to recover (Royal College of Physicians, 2020).

In the same time window as the 'local' effect (260-444ms), we conducted the average voltage interaction that failed to provide evidence for a significant interaction, although the observed signal showed the same pattern as the 'Global' effect that was observed in healthy participants.

The aim of this thesis is to propose a paradigm that can provide evidence for the use of strategic semantic expectations when processing language stimuli. We conducted four separate experiments including behavioural and electrophysiological measures, where we identify a relatedness proportion manipulation (Hutchison, 2007) that can distinguish between strategic and non-strategic semantic expectations. We provided DOC patient cases that aimed at exploring the potential results that could be observed in patients and how these results would differ from those of healthy participants. Future studies should assess this auditory tool in larger samples of DOC patients, to investigate how these effects behave across a larger group of patients; how the results differ between diagnostic labels (i.e., VS/UWS, MCS-, MCS+, etc.); and whether the presence or absence of these effects has any implications for patients' prognosis.

As previously mentioned, this relatedness proportion manipulation (Hutchison, 2007) follows the same underlying structure or rationale as the local-global paradigm, which is broadly used in DOC research as it differentiates between automatic and consciously controlled processing, by measuring the involvement of local and global expectations when individuals process auditory regularities (Sergent et al., 2017; Faugeras et al., 2012; Bekinschtein et al., 2009; El Karoui et al. 2015). Both paradigms create local and global contexts that encourage the recruitment of expectancy mechanisms at different time-scales and involving different levels of complexity; the local context enables the use of automatic or non-strategic expectations, whereas the global context entails engagement in controlled strategic expectations (Bekinschtein et al., 2009; Hutchison 2007), see Figure 1.1.However, one difference is that the local-global paradigm establishes the global context within each block, so individuals can only know the global regularity once they have listened a few trials, while this relatedness proportion paradigm creates the context across the task where the prime carries the global context cue, so only one trial is needed to know to which global context it belongs (Bekinschtein et al., 2009; Hutchison, 2007). The experimental conditions in both paradigms are equivalent: 1) **H**igh **R**elated = **L**ocal **S**tandard – **G**lobal **S**tandard; 2) **H**igh **U**nrelated = **L**ocal **D**eviant – **G**lobal **D**eviant; 3) **L**ow **R**elated = **L**ocal **S**tandard – **G**lobal **D**eviant; 4) **L**ow **U**nrelated = **L**ocal **D**eviant – **G**lobal **S**tandard (Bekinschtein et al., 2009; Hutchison 2007). Moreover, the data analyses in both paradigms are mathematically equivalent, where both local effects are calculated as follows: 1) Local-global paradigm = ((LDGD + LDGS)/2 – (LSGS + LSGD)/2); 2) Prime validity paradigm = ((HU + LU)/2 – (HR + LR)/2). Regarding both global effects, the calculations are as follows: 1) Local-global paradigm= (LDGD – LSGS) – (LDGS – LSGD) = (LDGD + LSGD)/2 – (LSGS + LDGS)/2; 2) Prime validity paradigm = (HU – HR) – (LU – LR) (Bekinschtein et al., 2009; Hutchison 2007). The similarity between both paradigms, will allow researchers and clinicians to use them as complementary measures to detect controlled strategic processing using different stimuli, and due to the similarities mentioned above, both results could be directly compared. Future research on the diagnosis of DOC could combine both paradigms to assess patients' level of awareness, as both paradigms measure the involvement in strategic conscious expectations in two different modalities (i.e., auditory tones and semantic stimuli).

A previous study proposed a two-stage profile (Rohaut et al., 2015) where the N400 effect reflects the processing of semantic stimuli that is non-conscious, followed by an LPC indicating conscious access to words. The authors conducted a semantic priming task, where they observed both the N400 effect and an LPC in healthy controls, both effects showing more extreme voltage values for incongruent (i.e., unrelated) targets relative to congruent (i.e., related) targets. Moreover, in the DOC group the authors observed N400 effects, whereas the results in the same group failed to show an LPC effect, although this effect was present in the MCS subgroup. On a single subject level, both the N400 and LPC effects were detected in 8 out of 19 healthy controls; and both the N400 and LPC effects were observed in 6 out of 29 patients (5 MCS, 1 VS/UWS) (Rohaut et al., 2015). The LPC was not detected in VS/UWS patients (Rohaut et al., 2015), which provides evidence that some conscious involvement is required to detect this effect; however, the presence or absence of this effect is based on a semantic priming task, which by itself is not sufficient to accredit strategic involvement (Hutchison, 2007; Neely & Keefe, 1989). The 'global' effect (i.e., prime validity interaction) that we are proposing has a direct involvement with the global context in which stimuli are presented, thus, requiring engagement in conscious expectations, in the same manner as the P3b requires conscious processing to be detected in the local-global paradigm (Bekinschtein et al., 2009). Moreover, our global effect is supported by individuals self-report about the use of a conscious strategy, and higher confidence ratings in high validity contexts that suggest the use of global expectations.

The detection of N400 effects in single-subject has been described as challenging (Kallionpää et al., 2019; Rohaut et al., 2015; Cruse et al, 2014), due to inconsistent methods, therefore, it is recommended to use more than one method to analyse these effects (Kallionpää et al., 2019). For example, Kallionpää and colleagues (2019) compared several possible methods to analyse single-subjects' data that included visual inspection, average in a time window ANOVA, cluster-based non-parametric testing, Bayesian method, and t-CWT. Regarding the analyses that we are here proposing, our local effect is estimated using a cluster-based non parametric testing that should be visually inspected; moreover, our global effect is obtained by taking the values from significant clusters from the cluster-based non parametric testing, and then conduct an average in a time window ANOVA analysis, and should also include visual inspection. Therefore, future studies should investigate different analyses approaches with DOC patients' data, to produce the best possible sensitivity in single-subjects to detect conscious language processing. Thus, avoiding both false positives (e.g., detect the ERP interaction but the patient is not conscious) that can give false hope for families and mislead the patient's future treatment direction; and false negatives (e.g., fail to detect the ERP interaction but the patient is conscious) that can undermine a patient’s potential for rehabilitation (Faugeras et al., 2012; Racine et al., 2010; Parikh, Mathai, Parikh, Sekhar, & Thomas, 2008).

Final Conclusions

The present thesis provides evidence for the use of strategic semantic expectations in a relatedness proportion paradigm in a semantic priming task. We here report behavioural RPEs in healthy participants from three behavioural experiments, where we detect greater semantic priming effects in targets that are presented in a high validity context, relative to targets presented in a low validity context. The presence of semantic priming effects indicate the use of expectations that are generated from a local context (i.e., related and unrelated targets), whereas RPEs suggest the use of top-down expectations that are generated from the global context in which words are presented (i.e., prime validity manipulation). Moreover, as a novel finding from this thesis we found that the RPEs are modulated by the use of strategy throughout the task, as individuals that report using the primes strategically show RPEs, in contrast to individuals that do not report using the prime strategically and thus failed to show RPEs.

Regarding the neural markers of strategic semantic expectations in this visual paradigm we report an early local ERP effect around 250ms that shows a predictive pattern and is similar to the MMN elicited in the local-global paradigm. This local effect is followed by a late global ERP effect around 350ms that interacts with the global context, as it involves the use of the prime validity, however, this effect shows an apredictive signal. Furthermore, we adapted the task to meet DOC patients' abilities, so we design an auditory manipulation that only includes a mental task, instead of measuring pronunciation responses as in the previous experiment. In a group of healthy participants, we report an early local ERP effect around 350ms that shows a predictive signal, similar to the N400 effect; followed by a global ERP effect between 300-800ms that involves the use of the global context and shows a predictive signal.

Although both ERP studies include the same paradigm, the visual experiment has a motor task, and the global ERP effect shows an apredictive signal that has the same pattern as behavioural responses; while the auditory experiment only has a mental task, and the global ERP effect shows a predictive signal. The difference in the direction of the ERP signal between both studies is a novel finding from this thesis, as the apredictive signal can be explained as neural signals being routed towards behaviour, whereas the predictive signal in the auditory task may reveal purer measures of cognitive processes (i.e., prediction error signal) due to the absence of a motor task. Furthermore, we present two patient cases (1 VS/UWS, 1 MCS) who performed the auditory task, where the VS/UWS patient fails to show both the local and global effects; while the MCS patient shows a significant local effect and failed to detect a significant global effect, however, the ERP signal follows the same pattern as healthy participants. This auditory paradigm has potential to help identifying strategic expectations in language processing in the context of DOC patients; however, due to the lack of single-subject interactions in healthy controls, this paradigm requires more experimental testing before being considered as a clinically viable tool to measure strategic semantic processing. This task follows the same structure as the local-global paradigm as it differentiates non-strategic processing (i.e., semantic local effect, MMN) from strategic processing (i.e., semantic global effect, P3b) in the domain of language comprehension; and the presence of conscious strategic processing can be considered as a sign of consciousness.

REFERENCES

1. Andrews, K., Murphy, L., Munday, R., & Littlewood, C. (1996). Misdiagnosis of the vegetative state: retrospective study in a rehabilitation unit. *Bmj*, 313(7048), 13-16.
2. Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production?. *Journal of Memory and Language*, 24(1), 89-106.
3. Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. Journal of Experimental Psychology: *Learning, memory, and cognition*, 12(3), 336.
4. Barascud, N., Pearce, M. T., Griffiths, T. D., Friston, K. J., & Chait, M. (2016). Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proceedings of the National Academy of Sciences*, 113(5), E616-E625.
5. Bauer, M., Stenner, M. P., Friston, K. J., & Dolan, R. J. (2014). Attentional modulation of alpha/beta and gamma oscillations reflect functionally distinct processes. *Journal of Neuroscience*, 34(48), 16117-16125
6. Beaumont, J. G., & Kenealy, P. M. (2005). Incidence and prevalence of the vegetative and minimally conscious states. *Neuropsychological rehabilitation*, 15(3-4), 184-189.
7. Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & cognition*, 8(6), 493-512.

8. 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.

9. Beukema, S., Gonzalez-Lara, L. E., Finoia, P., Kamau, E., Allanson, J., Chennu, S., ... & Cruse, D. (2016). A hierarchy of event-related potential markers of auditory processing in disorders of consciousness. *NeuroImage*: Clinical, 12, 359-371.

10. Blume, C., Del Giudice, R., Wislowska, M., Lechinger, J., & Schabus, M. (2015). Across the consciousness continuum—from unresponsive wakefulness to sleep. *Frontiers in human neuroscience*, 9, 105

11. Bodien, Y. G., Giacino, J. T., & Edlow, B. L. (2017). Functional MRI motor imagery tasks to detect command following in traumatic disorders of consciousness. *Frontiers in neurology*, 8, 688.

12. Bonhage, C. E., Mueller, J. L., Friederici, A. D., & Fiebach, C. J. (2015). Combined eye tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. *Cortex*, 68, 33-47.

13. Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Towards a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in psychology*, 10, 298.

14. Bostanov, V., & Kotchoubey, B. (2006). The t-CWT: a new ERP detection and quantification method based on the continuous wavelet transform and Student's t-statistics. *Clinical Neurophysiology*, 117(12), 2627-2644.

15. Boudewyn, M. A., Long, D. L., & Swaab, T. Y. (2015). Graded expectations: Predictive processing and the adjustment of expectations during spoken language comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, 15(3), 607-624.

16. Bowman, H., Filetti, M., Wyble, B., & Olivers, C. (2013). Attention is more than prediction precision [Commentary on target article]. *Behavioral and Brain Sciences*, 36(3), 206-208.

17. Brainard, D. H. (1997) The Psychophysics Toolbox, Spatial Vision 10:433-436.

18. Brookes, M. (1997). Voicebox: Speech processing toolbox for matlab. Software, available [Mar. 2011] from www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html, 47.

19. Brookes, M. J., Tewarie, P. K., Hunt, B. A., Robson, S. E., Gascoyne, L. E., Liddle, E. B., ... & Morris, P. G. (2016). A multi-layer network approach to MEG connectivity analysis. *Neuroimage*, 132, 425-438.

20. Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, 93, 203-216.

21. Brown, C. A., Seymour, B., Boyle, Y., El-Deredy, W., & Jones, A. K. (2008). Modulation of pain ratings by expectation and uncertainty: Behavioral characteristics and anticipatory neural correlates. *Pain*, 135(3), 240-250.

22. Brown, C. M., Hagoort, P., & Chwilla, D. J. (2000). An event-related brain potential analysis of visual word priming effects. *Brain and language*, 72(2), 158-190.

23. Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature reviews neuroscience*, 14(5), 365-376.

24. Cavanna, A. E., Shah, S., Eddy, C. M., Williams, A., & Rickards, H. (2011). Consciousness: a neurological perspective. *Behavioural neurology*, 24(1), 107-116

25. Chennu, S., & Bekinschtein, T. A. (2012). Arousal modulates auditory attention and awareness: insights from sleep, sedation, and disorders of consciousness. *Frontiers in psychology*, 3, 65.

26. Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibánez, A., ... & Bekinschtein, T. A. (2013). Expectation and attention in hierarchical auditory prediction. *Journal of Neuroscience*, 33(27), 11194-11205.

27. Childs, N. L., Mercer, W. N., & Childs, H. W. (1993). Accuracy of diagnosis of persistent vegetative state. *Neurology*, 43(8), 1465-1465.

28. Chwilla, D. J., Hagoort, P., & Brown, C. M. (1998). The mechanism underlying backward priming in a lexical decision task: Spreading activation versus semantic matching. *The Quarterly Journal of Experimental Psychology Section A*, 51(3), 531-560.

29. Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, 36(3), 181-204.

30. Cohen, M. X. (2014). Analyzing neural time series data: theory and practice. *MIT Press.*

31. Coleman, M. R., Davis, M. H., Rodd, J. M., Robson, T., Ali, A., Owen, A. M., & Pickard, J. D. (2009). Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. *Brain*, 132(9), 2541-2552.

32. Coleman, M. R., Rodd, J. M., Davis, M. H., Johnsrude, I. S., Menon, D. K., Pickard, J. D., & Owen, A. M. (2007). Do vegetative patients retain aspects of language comprehension? Evidence from fMRI. *Brain*, 130(10), 2494-2507.

33. Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, 82(6), 407.

34. Cruse, D., & Owen, A. M. (2010). Consciousness revealed: new insights into the vegetative and minimally conscious states. *Current Opinion in Neurology*, 23(6), 656-660.

35. Cruse, D., Beukema, S., Chennu, S., Malins, J. G., Owen, A. M., & McRae, K. (2014). The reliability of the N400 in single subjects: implications for patients with disorders of consciousness. *NeuroImage: Clinical*, 4, 788-799.

36. 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. *The Lancet*, 378(9809), 2088-2094.

37. Cruse, D., Chennu, S., Fernández-Espejo, D., Payne, W. L., Young, G. B., & Owen, A. M. (2012). Detecting awareness in the vegetative state: electroencephalographic evidence for attempted movements to command. *PLoS One*, 7(11), e49933.

38. Cruse, D., Monti, M. M., & Owen, A. M. (2011). Neuroimaging in disorders of consciousness: contributions to diagnosis and prognosis. *Future Neurology*, 6(2), 291-299.

39. Csikszentmihalyi, M., & Nakamura, J. (2010). Effortless attention in everyday life: A systematic phenomenology. *Effortless attention: A new perspective in the cognitive science of attention and action*, 179-189.

40. de Groot, A. M. (1984). Primed lexical decision: Combined effects of the proportion of related prime-target pairs and the stimulus-onset asynchrony of prime and target. *The Quarterly Journal of Experimental Psychology*, 36(2), 253-280.

41. De Houwer, J., Hermans, D., & Spruyt, A. (2001). Affective priming of pronunciation responses: Effects of target degradation. *Journal of experimental social psychology*, 37(1), 85-91.

42. De Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception?. *Trends in cognitive sciences*, 22(9), 764-779.

43. De Sousa, A. (2013). Towards an integrative theory of consciousness: part 1 (neurobiological and cognitive models). *Mens sana monographs*, 11(1), 100

44. Deacon, D., Hewitt, S., Yang, C. M., & Nagata, M. (2000). Event-related potential indices of semantic priming using masked and unmasked words: evidence that the N400 does not reflect a post-lexical process. *Cognitive Brain Research*, 9(2), 137-146.

45. Dehaene, S., & Christen, Y. (Eds.). (2011). Characterizing consciousness: from cognition to the clinic? *Springer Science & Business Media.*

46. DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature neuroscience,* 8(8), 1117-1121.

47. Delorme, A. & Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, 134: 9-21.

48. Dewan, M. C., Rattani, A., Gupta, S., Baticulon, R. E., Hung, Y. C., Punchak, M., ... & Rosenfeld, J. V. (2018). Estimating the global incidence of traumatic brain injury. *Journal of neurosurgery*, 130(4), 1080-1097

49. Drake, R., Vogl, A. W., & Mitchell, A. W. (2009). Gray's Anatomy for Students E-Book. *Elsevier Health Sciences.*

50. Dymowski, A. R., Owens, J. A., Ponsford, J. L., & Willmott, C. (2015). Speed of processing and strategic control of attention after traumatic brain injury. *Journal of clinical and experimental neuropsychology*, 37(10), 1024-1035.

51. El Karoui, I., King, J. R., Sitt, J., Meyniel, F., Van Gaal, S., Hasboun, D., ... & Cohen, L. (2015). Event-related potential, time-frequency, and functional connectivity facets of local and global auditory novelty processing: an intracranial study in humans. *Cerebral cortex*, 25(11), 4203-4212.

52. Fargier, R., Bürki, A., Pinet, S., Alario, F. X., & Laganaro, M. (2018). Word onset phonetic properties and motor artifacts in speech production EEG recordings. *Psychophysiology*, 55(2), e12982.

53. Faugeras, F., Rohaut, B., Weiss, N., Bekinschtein, T., Galanaud, D., Puybasset, L., ... & Naccache, L. (2012). Event related potentials elicited by violations of auditory regularities in patients with impaired consciousness. *Neuropsychologia*, 50(3), 403-418.

54. Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

55. Fernández-Espejo, D., & Owen, A. M. (2013). Detecting awareness after severe brain injury. *Nature Reviews Neuroscience*, 14(11), 801-809.

56. Fins, J. J. (2013). Disorders of consciousness and disordered care: families, caregivers, and narratives of necessity. *Archives of Physical Medicine and Rehabilitation*, 94(10), 1934-1939.

57. Fitz, H., & Chang, F. (2019). Language ERPs reflect learning through prediction error propagation. *Cognitive Psychology*, 111, 15-52.

58. Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society B: Biological sciences,* 360(1456), 815-836.

59. Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature reviews neuroscience*, 11(2), 127-138.

60. Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1211-1221.

61. Gazzaniga, M.S., Ivry, R.B. & Mangun, G.R. (2009). Cognitive Neuroscience: The Biology of the Mind (3rd ed.). *New York: W.W. Norton.*

62. Giacino, J. T., Ashwal, S., Childs, N., Cranford, R., Jennett, B., Katz, D. I., ... & Zasler, N. D. (2002). The minimally conscious state definition and diagnostic criteria. *Neurology*, 58(3), 349-353.

63. Giacino, J. T., Fins, J. J., Laureys, S., & Schiff, N. D. (2014). Disorders of consciousness after acquired brain injury: the state of the science. *Nature Reviews Neurology*, 10(2), 99.

64. Gibson, R. M., Fernández-Espejo, D., Gonzalez-Lara, L. E., Kwan, B. Y., Lee, D. H., Owen, A. M., & Cruse, D. (2014). Multiple tasks and neuroimaging modalities increase the likelihood of detecting covert awareness in patients with disorders of consciousness. *Frontiers in Human Neuroscience*, 8, 950.

65. Gill-Thwaites H, Munday R. (2004) The Sensory Modality Assessment and Rehabilitation Technique (SMART): a valid and reliable assessment for vegetative state and minimally conscious state patients*. Brain Inj*;18(12):1255–69.

66. Goldfine, A. M., & Schiff, N. D. (2011). Consciousness: its neurobiology and the major classes of impairment. *Neurologic clinics*, 29(4), 723-737.

67. Graham, M., Weijer, C., Cruse, D., Fernandez-Espejo, D., Gofton, T., Gonzalez-Lara, L. E., ... & Speechley, K. N. (2015). An ethics of welfare for patients diagnosed as vegetative with covert awareness. *AJOB Neuroscience*, 6(2), 31-41.

68. Griffiths, P. (2006) What is brain injury? Key facts about the initial stages of recovery. In Parker, J. and Price, D.J. (Eds.), Good practice in brain injury case management (pp. 31-47). *Jessica Kingsley Publishers*.

69. Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology*, 43(1), 21-30.

70. Gulan, T., & Valerjev, P. (2010). Semantic and related types of priming as a context in word recognition. *Review of psychology,* 17(1), 53-58.

71. Harrison, A. H., & Connolly, J. F. (2013). Finding a way in: a review and practical evaluation of fMRI and EEG for detection and assessment in disorders of consciousness. *Neuroscience & Biobehavioral Reviews*, 37(8), 1403-1419.

72. Heilbron, M., & Chait, M. (2018). Great expectations: is there evidence for predictive coding in auditory cortex?. *Neuroscience*, 389, 54-73.

73. Hertrich, I., Dietrich, S., & Ackermann, H. (2016). The role of the supplementary motor area for speech and language processing. *Neuroscience & Biobehavioral Reviews,* 68, 602-610.

74. Hill, H., Strube, M., Roesch-Ely, D., & Weisbrod, M. (2002). Automatic vs. controlled processes in semantic priming—differentiation by event-related potentials. *International Journal of Psychophysiology*, 44(3), 197-218.

75. Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in psychology*, 3, 96.

76. Hohwy, J. (2017). Priors in perception: Top-down modulation, Bayesian perceptual learning rate, and prediction error minimization. *Consciousness and Cognition*, 47, 75-85.

77. Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and language*, 35(1), 66-85.

78. Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and cognitive processes*, 5(4), 281-312.

79. Höller, Y., Bergmann, J., Thomschewski, A., Kronbichler, M., Höller, P., Crone, J. S., ... & Trinka, E. (2013). Comparison of EEG-features and classification methods for motor imagery in patients with disorders of consciousness. *PloS one*, 8(11), e80479.

80. Horga, G., & Maia, T. V. (2012). Conscious and unconscious processes in cognitive control: a theoretical perspective and a novel empirical approach. *Frontiers in human neuroscience*, 6, 199.

81. Horki, P., Bauernfeind, G., Klobassa, D. S., Pokorny, C., Pichler, G., Schippinger, W., & Müller-Putz, G. R. (2014). Detection of mental imagery and attempted movements in patients with disorders of consciousness using EEG. *Frontiers in Human Neuroscience*, 8, 1009.

82. Houses of Parliament (2015) Vegetative and Minimally Conscious States. *Postnote*; 489.

83. Huang, Y., & Rao, R. P. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(5), 580-593.

84. Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic bulletin & review*, 10(4), 785-813.

85. Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 645.

86. Hutchison, K. A., Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-Shikora, E. R., Tse, C. S., ... & Buchanan, E. (2013). The semantic priming project. *Behavior research methods,* 45(4), 1099-1114.

87. Hutchison, K. A., Neely, J. H., & Johnson, J. D. (2001). With great expectations, can two" wrongs" prime a" right"?. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 27(6), 1451.

88. Hyder, A. A., Wunderlich, C. A., Puvanachandra, P., Gururaj, G., & Kobusingye, O. C. (2007). The impact of traumatic brain injuries: a global perspective. *NeuroRehabilitation*, 22(5), 341-353.

89. JASP Team. (2017). JASP (Version 0.8. 3.1) [Computer software].

90. JASP Team. (2018). JASP (Version 0.9. 0.1) [Computer software].

91. JASP Team. (2019). JASP (Version 0.9.2.0) [Computer software].

92. Jennett, B., & Plum, F. (1972). Persistent vegetative state after brain damage: a syndrome in search of a name. *The Lancet*, 299(7753), 734-737.

93. Kahan, T. A., Neely, J. H., & Forsythe, W. J. (1999). Dissociated backward priming effects in lexical decision and pronunciation tasks. *Psychonomic Bulletin & Review*, 6(1), 105-110.

94. Kallionpää, R. E., Pesonen, H., Scheinin, A., Sandman, N., Laitio, R., Scheinin, H., ... & Valli, K. (2019). Single-subject analysis of N400 event-related potential component with five different methods. *International Journal of Psychophysiology*, 144, 14-24.

95. Kallionpää, R. E., Scheinin, A., Kallionpää, R. A., Sandman, N., Kallioinen, M., Laitio, R., ... & Scheinin, H. (2018). Spoken words are processed during dexmedetomidine-induced unresponsiveness. *British Journal of Anaesthesia*, 121(1), 270-280.

96. Kalmar, K., & Giacino, J. T. (2005). The JFK coma recovery scale—revised. Neuropsychological rehabilitation, 15(3-4), 454-460.

97. Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working memory capacity. *Journal of Experimental Psychology*: General, 130, 169 –183.

98. Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132, 47–70.

99. Keefe, D. E., & Neely, J. H. (1990). Semantic priming in the pronunciation task: The role of prospective prime-generated expectancies. *Memory & cognition*, 18(3), 289-298.

100. Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language,* 47(1), 145-171.

101. Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: Further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research*, 13(1), 27-39.

102. Kiefer, M., & Brendel, D. (2006). Attentional modulation of unconscious “automatic” processes: Evidence from event-related potentials in a masked priming paradigm. *Journal of cognitive neuroscience*, 18(2), 184-198.

103. King, J. R., Gramfort, A., Schurger, A., Naccache, L., & Dehaene, S. (2014). Two distinct dynamic modes subtend the detection of unexpected sounds. *PloS one,* 9(1), e85791.

104. Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3?.

105. Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 26(2), 441.

106. Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, 27(12), 712-719.

107. Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, 17(5), 307

108. Koivisto, M., & Revonsuo, A. (2001). Cognitive representations underlying the N400 priming effect. *Cognitive Brain Research*, 12(3), 487-490.

109. Kok, P., Rahnev, D., Jehee, J. F., Lau, H. C., & De Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral cortex*, 22(9), 2197-2206.

110. Koriat, A. (1981). Semantic facilitation in lexical decision as a function of prime-target association. *Memory & Cognition*, 9(6), 587-598.

111. Kotchoubey, B., & Pavlov, Y. G. (2018). A systematic review and meta-analysis of the relationship between brain data and the outcome in disorders of consciousness. *Frontiers in neurology*, 9, 315.

112. Kotchoubey, B., Lang, S., Mezger, G., Schmalohr, D., Schneck, M., Semmler, A., ... & Birbaumer, N. (2005). Information processing in severe disorders of consciousness: vegetative state and minimally conscious state. *Clinical neurophysiology*, 116(10), 2441-2453.

113. Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension?. *Language, cognition and neuroscience*, 31(1), 32-59.

114. Kuperberg, G. R., Delaney-Busch, N., Fanucci, K., & Blackford, T. (2018). Priming production: Neural evidence for enhanced automatic semantic activity preceding language production in schizophrenia. *NeuroImage: Clinical*, 18, 74-85.

115. Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology*, 62, 621-647.

116. Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.

117. Lapatki, B. G., Stegeman, D. F., & Jonas, I. E. (2003). A surface EMG electrode for the simultaneous observation of multiple facial muscles. *Journal of neuroscience methods*, 123(2), 117-128.

118. Lau, E. F., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2013). Automatic semantic facilitation in anterior temporal cortex revealed through multimodal neuroimaging. *Journal of Neuroscience*, 33(43), 17174-17181.

119. Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of cognitive neuroscience*, 25(3), 484-502.

120. Laureys S, Celesia GG, Cohadon F, et al. (2010) Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. *BMC Med 2010*; 8: 68.

121. Laureys, S., & Schiff, N. D. (2012). Coma and consciousness: paradigms (re) framed by neuroimaging. *Neuroimage*, 61(2), 478-491.

122. Laureys, S., Boly, M., Moonen, G., & Maquet, P. (2009). Two dimensions of consciousness: arousal and awareness. *Encycl Neurosci*, 2, 1133-1142.

123. Laureys, S., Owen, A. M., & Schiff, N. D. (2004). Brain function in coma, vegetative state, and related disorders. *The Lancet Neurology*, 3(9), 537-546.

124. Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, 68, 155-168.

125. Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. Frontiers in psychology, 6, 1171.

126. Luck, S. J. (2014). An introduction to the event-related potential technique. *MIT press.*

127. Lupker, S. J. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning and Verbal Behavior*, 23(6), 709-733.

128. Maniscalco, B., Lee, J. L., Abry, P., Lin, A., Holroyd, T., & He, B. J. (2018). Neural integration of stimulus history underlies prediction for naturalistically evolving sequences. *Journal of Neuroscience*, 38(6), 1541-1557.

129. Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, 164(1), 177-190.

130. Mathworks, I. (2016). MATLAB: R2016b. Mathworks Inc, Natick.

131. Mathworks, I. (2018). MATLAB: R2018a. Mathworks Inc, Natick.

132. Mathworks, I. (2020). MATLAB: R2020a. Mathworks Inc, Natick.

133. Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology,* 48(2), 229-240.

134. Monti, M. M., Laureys, S., & Owen, A. M. (2010). The vegetative state. *Bmj*, 341, c3765.

135. Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., ... & Laureys, S. (2010). Willful modulation of brain activity in disorders of consciousness. *New England Journal of Medicine*, 362(7), 579-589.

136. Moritz, S., & Graf, P. (2006). Performance of depressive and healthy participants on an affective priming task using word pronunciation. *German Journal of Psychiatry*, 9(2006), 1-9.

137. Moritz, S., Mersmann, K., Kloss, M., Jacobsen, D., Andresen, B., Krausz, M., ... & Naber, D. (2001). Enhanced semantic priming in thought-disordered schizophrenic patients using a word pronunciation task. *Schizophrenia Research*, 48(2-3), 301-305.

138. Multi-Society Task Force on PVS (1994) Medical aspects of the persistent vegetative state (1). *N Engl J Med*. 330(21):1499-1508.

139. Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clinical neurophysiology*, 115(1), 140-144.

140. National Institute for Health and Care Excellence. (2014). Head injury: assessment and early management. London: NICE Guideline (CG176)

141. Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition*, 4(5), 648-654.

142. Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of experimental psychology: general*, 106(3), 226.

143. Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic processes in reading: Visual word recognition*, 11(1), 264-336.

144. Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective-retrospective processing theory. *In Psychology of learning and motivation* (Vol. 24, pp. 207-248). Academic Press.

145. Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(6), 1003.

146. Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). The University of South Florida word association, rhyme and fragment norms. Retrieved February, 5, 2004.

147. Nguyen, R., Fiest, K. M., McChesney, J., Kwon, C. S., Jette, N., Frolkis, A. D., ... & Pringsheim, T. (2016). The international incidence of traumatic brain injury: a systematic review and meta-analysis. *Canadian journal of neurological sciences*, 43(6), 774-785.

148. Noesselt, T., Shah, N. J., & Jäncke, L. (2003). Top-down and bottom-up modulation of language related areas–an fMRI study. *BMC neuroscience*, 4(1), 13.

149. Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: fully automated statistical thresholding for EEG artifact rejection. *Journal of neuroscience methods*, 192(1), 152-162.

150. Nourski, K. V., Steinschneider, M., Rhone, A. E., Kawasaki, H., Howard, M. A., & Banks, M. I. (2018). Auditory predictive coding across awareness states under anesthesia: an intracranial electrophysiology study. *Journal of Neuroscience*, 38(39), 8441-8452.

151. Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 2011, 1

152. Owen, A. M. (2008). Disorders of consciousness. *Annals of the New York Academy of Sciences*, 1124(1), 225-238.

153. Owen, A. M., & Coleman, M. R. (2008). Functional neuroimaging of the vegetative state. *Nature Reviews Neuroscience*, 9(3), 235-243.

154. Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science*, 313(5792), 1402-1402.

155. Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian journal of ophthalmology*, 56(1), 45.

156. Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. (1998). Does pizza prime coin? Perceptual priming in lexical decision and pronunciation. *Journal of Memory and Language,* 38(4), 401-418.

157. Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision* 10:437-442.

158. Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition,* 62(2), 223-240.

159. Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological research*, 66(3), 180-194.

160. Perrin, F., Bastuji, H., & Garcia-Larrea, L. (2002). Detection of verbal discordances during sleep. *Neuroreport*, 13(10), 1345-1349.

161. Perrin, F., Schnakers, C., Schabus, M., Degueldre, C., Goldman, S., Brédart, S., ... & Maquet, P. (2006). Brain response to one's own name in vegetative state, minimally conscious state, and locked-in syndrome. *Archives of neurology*, 63(4), 562-569.

162. Peterson, A., Cruse, D., Naci, L., Weijer, C., & Owen, A. M. (2015). Risk, diagnostic error, and the clinical science of consciousness. *NeuroImage: Clinical*, 7, 588-597.

163. Peterson, R. R., & Simpson, G. B. (1989). Effect of backward priming on word recognition in single-word and sentence contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 15(6), 1020.

164. Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension?. *Trends in cognitive sciences*, 11(3), 105-110.

165. Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In Proceedings of the 17th annual conference of the cognitive science society (Vol. 17, No. 2, pp. 37-42).

166. Popov, T., Oostenveld, R., & Schoffelen, J. M. (2018). FieldTrip made easy: An analysis protocol for group analysis of the auditory steady state brain response in time, frequency, and space. *Frontiers in neuroscience*, 12, 711

167. Posner, J. B., & Plum, F. (2007). Plum and Posner's diagnosis of stupor and coma (Vol. 71).

168. Posner, J. B., Plum, F., Saper, C. B., & Schiff, N. (2007). Plum and Posner's diagnosis of stupor and coma (Vol. 71). OUP USA.

169. Posner, M. I., Snyder, C. R. R., & Solso, R. L. (1975). Theories in information processing.

170. Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral science*, 12(5), 410-430.

171. Rabovsky, M., & McRae, K. (2014). Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Cognition*, 132(1), 68-89.

172. Racine, E., Rodrigue, C., Bernat, J. L., Riopelle, R., & Shemie, S. D. (2010). Observations on the ethical and social aspects of disorders of consciousness. *Canadian journal of neurological sciences*, 37(6), 758-768.

173. Rämä, P., Relander-Syrjänen, K., Öhman, J., Laakso, A., Näätänen, R., & Kujala, T. (2010). Semantic processing in comatose patients with intact temporal lobes as reflected by the N400 event-related potential. *Neuroscience letters*, 474(2), 88-92.

174. 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.

175. Robinson, S. (1999) Functional neuroimaging by synthetic aperture magnetometry (SAM). *Recent advances in biomagnetism*

176. Rohaut, B., & Naccache, L. (2017). Disentangling conscious from unconscious cognitive processing with event-related EEG potentials. *Revue Neurologique*, 173(7-8), 521-528.

177. Rohaut, B., Eliseyev, A., & Claassen, J. (2019). Uncovering consciousness in unresponsive ICU patients: technical, medical and ethical considerations. *Critical Care*, 23(1), 78.

178. Rohaut, B., Faugeras, F., Chausson, N., King, J. R., El Karoui, I., Cohen, L., & Naccache, L. (2015). Probing ERP correlates of verbal semantic processing in patients with impaired consciousness. *Neuropsychologia*, 66, 279-292.

179. Rolke, B., Heil, M., Streb, J., & Hennighausen, E. (2001). Missed prime words within the attentional blink evoke an N400 semantic priming effect. *Psychophysiology*, 38(2), 165-174.

180. Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia,* 51(3), 437-447.

181. Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review*, 16(2), 225-237.

182. Royal College of Physicians (2013) Prolonged disorders of consciousness: National clinical guidelines. London, RCP.

183. Royal College of Physicians (2020) Prolonged disorders of consciousness following sudden onset brain injury: national clinical guidelines. Retrieved from: www.rcplondon.ac.uk/guidelines-policy/prolonged-disorders-consciousness-national-clinical-guidelines

184. Schnakers, C., Perrin, F., Schabus, M., Majerus, S., Ledoux, D., Damas, P., ... & Laureys, S. (2008). Voluntary brain processing in disorders of consciousness. *Neurology,* 71(20), 1614-1620.

185. Schnakers, C., Vanhaudenhuyse, A., Giacino, J., Ventura, M., Boly, M., Majerus, S., ... & Laureys, S. (2009). Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. *BMC neurology*, 9(1), 35.

186. Schoenle, P. W., & Witzke, W. (2004). How vegetative is the vegetative state? Preserved semantic processing in VS patients–evidence from N 400 event-related potentials. *NeuroRehabilitation,* 19(4), 329-334.

187. Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre-and postlexical loci of contextual effects on word recognition. *Memory & cognition*, 12(4), 315-328.

188. Sergent, C., Faugeras, F., Rohaut, B., Perrin, F., Valente, M., Tallon-Baudry, C., ... & Naccache, L. (2017). Multidimensional cognitive evaluation of patients with disorders of consciousness using EEG: a proof of concept study. *NeuroImage: Clinical,* 13, 455-469.

189. Seth, A. K. (2010). The grand challenge of consciousness. *Frontiers in psychology*, 1, 5.

190. Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends in cognitive sciences*, 12(8), 314-321.

191. Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming?. *Journal of Experimental Psychology: Learning, memory, and cognition*, 18(6), 1191.

192. Shiel, A., Horn, S. A., Wilson, B. A., Watson, M. J., Campbell, M. J., & McLellan, D. L. (2000). The Wessex Head Injury Matrix (WHIM) main scale: a preliminary report on a scale to assess and monitor patient recovery after severe head injury. *Clinical rehabilitation,* 14(4), 408-416.

193. Shomstein, S. (2012). Object-based attention: strategy versus automaticity. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(2), 163-169.

194. Sitt, J. D., King, J. R., El Karoui, I., Rohaut, B., Faugeras, F., Gramfort, A., ... & Naccache, L. (2014). Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. *Brain,* 137(8), 2258-2270.

195. Sokoliuk, R., Calzolari, S., & Cruse, D. (2019). Dissociable electrophysiological correlates of semantic access of motor and non-motor concepts. *Scientific reports,* 9(1), 1-14.

196. Sokoliuk, R., Mayhew, S. D., Aquino, K. M., Wilson, R., Brookes, M. J., Francis, S. T., ... & Mullinger, K. J. (2019). Two spatially distinct posterior alpha sources fulfil different functional roles in attention. *Journal of Neuroscience*, 39(36), 7183-7194.

197. Spieler, D. H., Balota, D. A., & Faust, M. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 461–479.

198. Steppacher, I., Eickhoff, S., Jordanov, T., Kaps, M., Witzke, W., & Kissler, J. (2013). N400 predicts recovery from disorders of consciousness. *Annals of Neurology*, 73(5), 594-602.

199. Stolz, J. A., Besner, D., & Carr, T. H. (2005). Implications of measures of reliability for theories of priming: Activity in semantic memory is inherently noisy and uncoordinated. *Visual Cognition*, 12(2), 284-336.

200. Strauss, M., Sitt, J. D., King, J. R., Elbaz, M., Azizi, L., Buiatti, M., ... & Dehaene, S. (2015). Disruption of hierarchical predictive coding during sleep. *Proceedings of the National Academy of Sciences*, 112(11), E1353-E1362.

201. Team, A. (2014). Audacity (R): free audio editor and recorder [Computer program].

202. Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness: a practical scale. *The Lancet,* 304(7872), 81-84.

203. Tennant, A. (2005). Admission to hospital following head injury in England: Incidence and socio-economic associations. *BMC public health*, 5(1), 21.

204. Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology*, 83(3), 382-392.

205. Ting, W. K. C., Perez Velazquez, J. L., & Cusimano, M. D. (2014). Eye movement measurement in diagnostic assessment of disorders of consciousness. *Frontiers in neurology*, 5, 137.

206. Turner-Strokes, L. (Ed.). (2003). Rehabilitation following acquired brain injury: national clinical guidelines. *Royal College of Physicians*

207. Van Berkum, J. J., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of cognitive neuroscience*, 11(6), 657-671.

208. Van Doorn, J., van den Bergh, D., Bohm, U., Dablander, F., Derks, K., Draws, T., ... & Ly, A. (2019). The JASP guidelines for conducting and reporting a Bayesian analysis.

209. Van Drongelen, W., Yuchtman, M., Van Veen, B. D., & Van Huffelen, A. C. (1996). A spatial filtering technique to detect and localize multiple sources in the brain. *Brain Topography,* 9(1), 39-49.

210. Van Erp, W. S., Lavrijsen, J. C. M., Van De Laar, F. A., Vos, P. E., Laureys, S., & Koopmans, R. T. C. M. (2014). The vegetative state/unresponsive wakefulness syndrome: a systematic review of prevalence studies. *European journal of neurology*, 21(11), 1361-1368.

211. van Gaal, S., De Lange, F. P., & Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Frontiers in human neuroscience*, 6, 121.

212. Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A*, 47(3), 631-650.

213. Van Veen, B. D., Van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on biomedical engineering*, 44(9), 867-880.

214. Vidal-Gran, C., Sokoliuk, R., Bowman, H., & Cruse, D. (2020) Strategic and non-strategic semantic expectations hierarchically modulate neural processing. *Eneuro*, 7(5).

215. Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... & Meerhoff, F. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic bulletin & review*, 25(1), 58-76.

216. Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature,* 203(4943), 380-384.

217. Weber, K., Lau, E. F., Stillerman, B., & Kuperberg, G. R. (2016). The yin and the yang of prediction: An fMRI study of semantic predictive processing. *PLoS One*, 11(3), e0148637.

218. Yap, M. J., Hutchison, K. A., & Tan, L. C. (2016). Individual differences in semantic priming performance: insights from the Semantic Priming Project. *Big Data in Cognitive Science: From Methods to Insights.*

219. Yates, P. J., Williams, W. H., Harris, A., Round, A., & Jenkins, R. (2006). An epidemiological study of head injuries in a UK population attending an emergency department. *Journal of Neurology, Neurosurgery & Psychiatry*, 77(5), 699-701

220. Yekutieli, D., & Benjamini, Y. (1999). Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical Planning and Inference,* 82(1-2), 171-196.

221. Ylinen, S., Huuskonen, M., Mikkola, K., Saure, E., Sinkkonen, T., & Paavilainen, P. (2016). Predictive coding of phonological rules in auditory cortex: A mismatch negativity study. *Brain and language*, 162, 72-80.

222. Young, G. B. (2009). Coma. *Annals of the New York Academy of Sciences*, 1157(1), 32-47.

223. Zeman, A. (2001). Consciousness. *Brain*, 124: 1263–1289.

224. Zylberberg, A., Slezak, D. F., Roelfsema, P. R., Dehaene, S., & Sigman, M. (2010). The brain's router: a cortical network model of serial processing in the primate brain. *PLoS Comput Biol*, 6(4), e1000765.