UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department: Padova Neuroscience Center

Ph.D. COURSE IN: Neuroscience

SERIES 36°

ROLE OF THE TEMPORO-PARIETAL JUNCTIONS AND PRE-STIMULUS DYNAMICS IN PREDICTIVE PROCESSING

Thesis written with the financial contribution of Name of the Financing Institution: San Camillo IRCCS s.r.l.

Coordinator: Ch.mo Prof. Antonino Vallesi

Supervisor: Ch.mo Prof. Antonino Vallesi

Co-Supervisor: Dr. Giorgio Arcara

P.h.D Student: Sara Lago

General Introduction: the role of the temporo-parietal junctions and pre-stimulus alpha in predictive processing

The temporoparietal junctions (TPJs) are areas of the cerebral cortex that roughly include bilateral ventral portions of the inferior parietal lobule (the supramarginal and angular gyri) and posterior sections of the superior temporal gyrus (Doricchi et al., 2022; Masina et al., 2022). Cytoarchitectonically, the TPJs encompass the PFop, PFt, PF, PFm, and PFcm, located in the supramarginal gyrus, and the PGp and PGa, located in the angular gyrus (Doricchi et al., 2022; Gillebert et al., 2013).

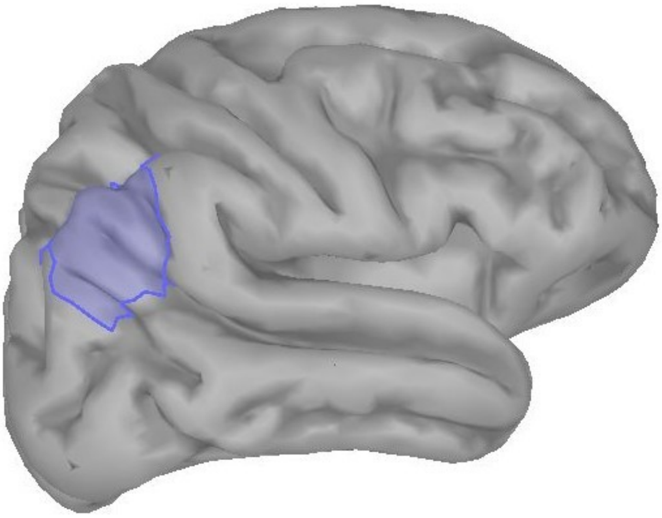


Figure 1. TPJ includes the supramarginal and angular gyri as well as the posterior part of the superior temporal gyrus.The figure highlights (in blue) the region encompassing the right TPJ.

These areas have been found to play a role in a variety of cognitive processes. For example, tasks that involve mentalizing, the ability to attribute mental states to one's self and others (Schurz et al., 2014, 2017), activate the right TPJ (rTPJ) - and, to a lesser extent, the left TPJ (lTPJ; e.g., Cabeza et al., 2012; Decety & Lamm, 2007; Igelström & Graziano, 2017).

Also body ownership (the perception of one's own body belonging to oneself) and sense of agency (the feeling of being in control of one's own actions) seem to depend in particular on the integrity of the rTPJ (Blanke et al., 2002; De Ridder et al., 2007), while the lTPJ seems to be active during episodic memory retrieval (Daselaar et al., 2006; Yonelinas et al., 2005) as well as during language processing: it is activated in response to lexical violations, non-words and words that are semantically unrelated with the previous context (Binder et al., 2005; Fiebach et al., 2002; Prince et al., 2007). In the language domain, lTPJ seems to play a role in integrating the individual's general world knowledge with the local discourse information (Menenti et al., 2008) and in pre-activatng linguistic information (Gastaldon et al., 2020). In addition, attentional tasks result in rTPJ activation, especially in response to unexpected but task-relevant stimuli (e.g., Doricchi et al., 2010; Indovina & Macaluso, 2007), leading to the hypothesis that rTPJ could play a key role in the reorienting of attention (Corbetta et al., 2008; Corbetta & Shulman, 2002). This hypothesis, known as the circuit-breaking theory, states that the dorsal attention network (DAN; Vossel et al., 2014) maintains visuospatial information relevant to the current task-defined goals, while the ventral attention network (VAN), that includes rTPJ, allows the reorientation of attention to behaviorally relevant and task-related but currently unattended stimuli; rTPJ would therefore be responsible for interrupting the activity of the DAN, resulting in the reorienting of attention to a new salient stimulus (Corbetta et al., 2008; Corbetta & Shulman, 2002). Moreover, rTPJ is not only part of the VAN: it is also encompassed in the default mode network (DMN), supporting mentalizing processes (Hughes et al., 2019; Schurz et al., 2014).

The involvement of the bilateral TPJs in such heterogeneous contexts has led to some domain-general conceptualizations of the role of these areas, in opposition to the domain-specific approaches outlined so far, that focus on each cognitive ability separately. Domain-general (or domain-independent) approaches strive to streamline neurobiological mechanisms that are common to multiple areas of interest such as perception, action and cognition, with its different domains (e.g., Corlett et al., 2022; Poldrack and Yarkoni, 2016).

For example, the nexus model (Carter & Huettel, 2013) aims at reconciling the roles of the TPJs in mentalizing and attention within a comprehensive theory. According to this model, the TPJs serve as central hubs, or "nexus", where lower-level functions such as attentional reorientation intersect with higher-order social-cognitive functions. The central concept of the nexus model is that the TPJs integrate information from various cognitive domains to create a social framework that supports decision-making. Geng and Vossel (2013) have in turn proposed an integrative view of the rTPJ specifically, which is based on the functional interpretation of the P300, an event-related potential (ERP) that arises from various neural sources. The P3b, a subcomponent of the P3, has been linked to TPJ and is considered a neural marker of contextual updating (Polich, 2007). According to this view, known as the contextual updating hypothesis, TPJ plays a role in updating internal models (representations) of the environment, which in turn generates appropriate expectations and behaviors (Geng & Vossel, 2013). In sum, these more general theories state that rTPJ in particular would act as a hub for integrating information from multiple domains and thus updating the internal models of the world.

Internal models are particularly important in predictive processing theories, which aim to provide comprehensive frameworks for understanding the neural processes involved in perception, cognition, and action (Clark, 2013; Friston, 2012). According to these frameworks, to deal with environmental changes, the brain predicts upcoming events through the implementation of internal models that spontaneously infer causal structures in the world (Clark, 2013). An internal model identifies causal regularities from a complex variety of sensory signals, extracting what is relevant to predict future stimuli, and detects mismatches between predictions (priors) and actual sensory inputs. The mismatch between a prediction and the actual sensory input generates a prediction error, which is used to update expectations and provide better predictions that are, in turn, tested against the actual input (Friston, 2019). Such predictive computations are believed to be carried out in a distributed way across the brain: a recent meta-analysis proposed that a set of brain areas, comprising also rTPJ among others, form a diffuse network involved in higher-level prediction generation and testing, that support both perception and action processes (Ficco et al., 2021; Siman-tov et al., 2019).

Interestingly, rTPJ is not only part of the putative prediction network, but it is also encompassed in both the VAN and the DMN (Hughes et al., 2019). This concurrency speaks in favor of a domain-general conceptualization of rTPJ as an area where multiple cognitive processes and information types converge and are integrated, but it also opens interesting perspectives on how rTPJ carries out cognitive computations in a seemingly ubiquitous way across domains, and on what consequences can arise in the case of a disruption of rTPJ or its disconnection from other areas or networks. For example, the involvement of rTPJ within the VAN suggests an intriguing role of this brain area in the interplay occurring between the VAN and the prediction network, in other words, between attention and prediction. Along with the well-established role of the VAN in the stimulus-driven changes in attentional focus, this network would be also activated in case of violation of expectations (Geng & Vossel, 2013), evidence that further links the VAN to the predictive processing framework. Specifically, Siman-Tov et al. (2019) argued that the VAN might be part, together with other brain structures, of a larger network subserving domain-general high-order prediction. The authors do not reject the idea that rTPJ is involved in attention, but they interpret it through the lens of predictive processing, which considers attention as an emergent property of the precision optimization mechanism that takes place during prediction. Precision is defined mathematically as the inverse variance of a prediction and represents its reliability (Friston, 2018), but it can also be associated with perceived stimuli or prediction errors (Walsh et al., 2020). In sum, the more precise a stimulus or a prediction, the more the individual will rely on it. In predictive terms, attending to a stimulus means representing and increasing the precision of sensory information (and prediction error) during the inferential process (Feldman & Friston, 2010), resulting in larger neural responses to attended vs. unattended stimuli (e.g., Jiang et al., 2013). In other words, attention depends in part on the predictions' precision (i.e., their reliability; Hohwy, 2012), and predictive processing proposes a constant interaction between attention and expectation (Walsh et al., 2020). In this vein, rTPJ, as a relevant hub of the VAN, would have a key role in estimating the precision of priors and sensory evidence. This hypothesis is not new, since attention has been already interpreted through the lens of predictive processing (Feldman & Friston, 2010; Hohwy, 2012). However, this scenario provides a sharpened interpretation of neural correlates of predictive mechanisms, and the overlap between the prediction network and the VAN substantiates the hypothesis that prediction and attention might be interdependent processes (Corbetta et al., 2008; Vossel et al., 2014).

Up to now we mainly discussed rTPJ; indeed, the evidence about the involvement of its left-hemisphere homologous in predictive processes is still limited. However, Doricchi et al. (2022) have proposed that the lTPJ has a role in prediction, by encoding both matches and mismatches between predicted and observed sensory, motor, or cognitive events, while rTPJ only encodes mismatches; in addition, these match/mismatch computations should follow the prediction error minimization principle. Following their line of reasoning, we can assume that the TPJs are involved in the prediction testing phase. Conversely, Siman-Tov et al. (2019) presume that their putative prediction network, encompassing rTPJ, is involved in both prediction generation and testing, a hypothesis supported by the probable interdependency between predictions' precision and attention (Hohwy, 2012; Walsh et al., 2020).

At this point, it should be noted that the majority of studies reviewed so far have mainly detected TPJ activation *after* predictions were disconfirmed. As a consequence, they only assume that the TPJs had a role in prediction, but they do not directly investigate it, and we can argue that the TPJs are involved in prediction violation, rather than in prediction *per se*. Here, therefore, some questions arise: are both left and right TPJ involved in prediction? Are they involved in prediction testing (i.e., checking whether the stimuli are or not in line with the predictions), as the majority of studies find, or also in prediction generation? If so, can the pre-stimulus state of the TPJs be representative of prediction generation? Answering these questions means investigating the TPJs' spontaneous activity during cognitive tasks, not only after target stimuli but also before them. We, therefore, have to direct our interest toward TPJs' state during the pre-stimulus interval, which could be representative of prediction generation, and, as such, might modulate the neural responses to subsequent stimuli. To this aim, we will use fast-resolution techniques such as MEG and EEG, and we will focus on neurophysiological dynamics that are known to be associated with predictive processes, such as pre-stimulus alpha oscillations (Alamia & VanRullen, 2019; Cao et al., 2017).

Aim of the dissertation

Broadly, the present dissertation aims at clarifying the role of the bilateral TPJs in neurological patients and healthy individuals, in domain-general and domain-specific cognition, with a specific interest in the language domain, which has been surprisingly overlooked during previous investigations on the predictive function of the TPJs. To investigate predictions, we will focus on the pre-stimulus interval, considered as a time interval when hypotheses about the upcoming stimuli are built by the individual on the basis of previous stimuli, the context, or prior knowledge.

We will first outline, in a systematic literature review, the role of the rTPJ in aberrant predictive processing as can be observed in neurological pathologies, a field in which evidence about rTPJ role is abundant. Even though our interest lies in the bilateral TPJs, we initially focus on rTPJ in Study 1 because of theoretical and empirical reasons. Dominant and unified theoretical accounts of the functional role of the TPJs concentrate mainly on attention, a fundamental cognitive function that plays a pivotal role in various other cognitive domains (Sani, 2023). This consideration is corroborated by long-standing neuropsychological evidence showing that up to 50% of stroke patients in the acute stage show signs of hemispatial neglect (Vallar & Calzolari, 2018), a disorder of attention that mainly follows lesions to the right hemisphere and, particularly, to rTPJ (Corbetta & Shulman, 2011). Probably for these reasons, rTPJ has been investigated more extensively with respect to its left-hemisphere counterpart, resulting in a more extended body of published studies: for example, when searching PubMed for articles on lTPJ, 867 results appear, while when searching for resources on rTPJ, 1226 entries are obtained[1].

After establishing the domain-general, predictive role of rTPJ, in Study 2 we will delve more in-depth into linguistic predictive dynamics, by investigating the unfolding of the neurophysiological correlates of linguistic and semantic prediction generation and violation. From this moment on, this work will mainly focus on prediction in the language domain and its relation with TPJs activity. Prediction is known to take place in language (e.g., Gastaldon et al., 2020; Nicenboim et al., 2020; Nieuwland et al., 2020), but it has been mostly investigated with violation paradigms, which only consider what happens after predictions have been (dis)confirmed (similarly to those studies on the predictive function of the TPJs, that detect activations in conditions of prediction violation; see above). Relatively few studies focused on the pre-stimulus interval and found that stronger expectations are associated with lower prestimulus alpha power. Traditionally, alpha oscillations have been conceptualized as a gate for information to be redirected to task-relevant brain regions, while task-irrelevant ones are inhibited (Jensen & Mazaheri, 2010), so that anticipatory increases in alpha power reflect a state of cortical inhibition, while reductions reflect cortical activation and facilitate the subsequent detection and processing of task-relevant information (Klimesch et al., 2007). Moreover, recent evidence specific for language processing points to a possible role of alpha oscillations during reading. For example, Jensen et al. (2021) highlighted a link between the timing of saccades while reading a text and alpha oscillations,

---

[1] Search algorithms: ((right [Title/Abstract]) AND (TPJ[Title/Abstract] OR temporal parietal junction[Title/Abstract] OR temporoparietal junction[Title/Abstract] OR temporal-parietal junction[Title/Abstract] OR temporo-parietal junction[Title/Abstract] OR temporo-parietal region[Title/Abstract])) OR (rTPJ[Title/Abstract]) for rTPJ and ((left [Title/Abstract]) AND (TPJ[Title/Abstract] OR temporal parietal junction[Title/Abstract] OR temporoparietal junction[Title/Abstract] OR temporal-parietal junction[Title/Abstract] OR temporo-parietal junction[Title/Abstract] OR temporo-parietal region[Title/Abstract])) OR (lTPJ[Title/Abstract]) for lTPJ. The research has been conducted on the 15.09.2023.

so that fixated and parafoveal words are represented at different phases of an alpha cycle. More specifically, onsets of saccades towards low frequency words within the text seem to be locked to the phase of alpha oscillations, suggesting that the alpha oscillations time the processing between the oculomotor and visual systems during reading, and that this coordination becomes more pronounced for demanding words (Pan et al., 2023). However, alpha power also fluctuates spontaneously, in the absence of experimental manipulations; and in the attention domain, spontaneously low prestimulus power is associated with better behavioral performance and with ERPs with shorter latencies and higher amplitudes. On the other hand, in the context of predictive processing, alpha oscillations in the pre-stimulus interval have been linked to top-down predictions (Alamia & Van Rullen, 2019; Cao et al., 2017), and especially to the concept of precision, the level of reliability of a prediction (see Section 1.2. in Study 2; Bauer et al., 2014; Sedley et al., 2016; Sherman et al., 2016). However, little is known about the role of precision and pre-stimulus alpha fluctuations in domains other than attention, for example, in language. To this aim, Study 2 investigates whether spontaneous fluctuations in prestimulus alpha power modulate language-related ERPs in a semantic congruence task. This allows us to better explore the role of pre-stimulus alpha power in predictive processing and to gain a deeper understanding of the link between pre- and post-stimulus neurophysiological correlates of predictive dynamics, i.e., of prediction generation and testing.

Finally, Study 3 will combine the investigation of pre-stimulus alpha with that of the role of bilateral TPJs during complex linguistic computations taking place during a metaphor comprehension task, a particular case of linguistic predictive processing (Vespignani et al., 2010). Using source-reconstructed MEG recordings, we selected areas from the language network in addition to left and right TPJ, with the aim of investigating the role of the bilateral TPJs in prediction generation and testing. More specifically, we wanted to answer the following questions: are both TPJs involved in linguistic prediction? Are they involved in both prediction generation and testing, as proposed by Siman-Tov et al. (2019), or only in prediction testing, as proposed by Doricchi et al. (2022)? If the TPJs are involved in prediction generation, can pre-stimulus alpha (associated with predictions' precision) modulate the subsequent brain responses to target stimuli? And finally, is the eventual modulation local, i.e., limited to the TPJs under investigation, or can pre-stimulus TPJ activity influence post-stimulus activations in other task-related areas? The results of this study can help shed some light on the possibly different roles of left and right TPJ during complex linguistic computations, and at what stage of predictive processing they come into play.

Study 1: Disconnection from prediction: A systematic review on the role of right temporoparietal junction in aberrant predictive processing

This study was published in Neuroscience and Biobehavioral Reviews: Masina, F., Pezzetta, R., Lago, S., Mantini, D., Scarpazza, C., & Arcara, G. (2022). Disconnection from prediction: A systematic review on the role of right temporoparietal junction in aberrant predictive processing. *Neuroscience & Biobehavioral Reviews*, *138*(January), 104713–104713. https://doi.org/10.1016/j.neubiorev.2022.104713

1. Introduction

The right temporoparietal junction (rTPJ) is a brain area that plays a critical role in the higher-order cognitive and motor functions that underlie human behavior. A lesion of this area or its disconnection from other brain structures may have several etiologies, ranging from trauma to neurodegenerative diseases. Among them, stroke (Campbell and Khatri, 2020) is the neurological condition for which the role of rTPJ has been more extensively investigated. Indeed, over half of all ischemic strokes occur in the middle cerebral artery territory (Ng et al., 2007), which supplies blood to the bilateral TPJ. The pattern of behavioral symptoms emerging from rTPJ lesions has fostered an extensive clinical interest in this brain structure. Lesions involving rTPJ are often associated with hemispatial neglect (Corbetta and Shulman, 2011), a disorder of attention that may follow right hemisphere stroke in up to 50 % of patients in the acute stage (Vallar and Calzolari, 2018) and related to a difficulty to orientate, report, or respond to stimuli located on the contralesional hemifield. Besides neglect, recent research effort has shown other functional consequences of rTPJ lesions in several neurological disorders (for extensive reviews, see Frucht et al., 2021; Pisella et al., 2011; Vallar, 2007), as spatial and visual impairments, and movement disorders. Thus, rTPJ potentially carries a high clinical significance and its deeper functional understanding could be extremely valuable for clinicians. The purpose of this systematic review is therefore to elucidate the involvement of rTPJ in human behavior. To this aim, we will investigate the functional role of rTPJ through the analysis of the functional consequences of its direct lesion or its disconnection from other regions in neurological disorders. We will initially discuss the role of rTPJ according to the most dominant theoretical accounts, and then introduce the predictive processing framework, an overarching and unifying theory that postulates that the core function of the brain is to minimize prediction errors concerning a generative model of the world (Pezzulo et al., 2021). In this vein, recently rTPJ has been conceived as a key hub of a putative prediction network (Siman-Tov et al., 2019). After presenting a systematic review of the literature linking rTPJ and neurological disorders, results will be interpreted in the light of a predictive processing framework. Implications and advantages of such interpretations will be highlighted, suggesting that impairments associated with this structure may cause aberrant and domain-general predictive mechanisms, rather than fragmented domain-specific cognitive deficits. Future endeavors for basic and clinical research will finally be proposed.

1.1. Dominant theoretical accounts of rTPJ in cognition

Anatomically, rTPJ refers to a composite portion of the cortex roughly encompassing the inferior parietal lobule and extending into the superior temporal gyrus. Albeit this definition can approximate the location of rTPJ, it does not account for its anatomical and cytoarchitectonic complexity (Bzdok et al., 2013; Caspers et al., 2006; Patel et al., 2019). Functionally, the activation of rTPJ has been associated with the involvement of a variety of cognitive functions, raising controversy concerning the role of this region (Krall et al., 2015). To date, there is little consensus regarding whether these functions are supported by specific mechanisms/modules, or they result from a common and overarching mechanism. Actually, there is evidence in favor of both functional specialization (Krall et al., 2015; Scholz et al., 2009) and functional overlap (Carter and Huettel, 2013; Corbetta et al., 2008; Geng and Vossel, 2013; Kubit and Jack, 2013; Wilterson et al., 2021). The functional specialization framework proposes a one-to-one mapping of cognitive mechanisms and brain structures (Scholz et al., 2009). It has been demonstrated that the posterior portion of rTPJ, which includes the angular gyrus and the terminal part of the superior temporal sulcus, plays a role in mentalizing, also known as Theory of Mind (ToM). Mentalizing refers to the ability to attribute mental states to one's self and others (Schurz et al., 2014, 2017) and represents a milestone in the development of social skills. Some theories posit that ToM is grounded in the capacity to construct inner models of someone else's beliefs (Koster-Hale and Saxe, 2013) that allow inferring the mental states of others (Frith et al., 2003). Recent studies have shown that mentalizing tasks involve cognitive and affective processes to different extents, but such socio-cognitive functions are hierarchically organized and mostly localized within rTPJ, activating different portions of the area (Schurz et al., 2017) or different networks encompassing rTPJ (Schurz et al., 2021). Along a posterior-anterior axis, evidence from a meta-analysis indicates that ToM shows higher activation probability in the posterior part of rTPJ whereas the most anterior portion has been associated with attention (Decety and Lamm, 2007). A large body of evidence shows the role of rTPJ as part of the ventral attentional control network (VAN) including the middle frontal gyrus and the inferior frontal gyrus. According to a dominant theory of rTPJ function, the VAN allows the reorientation of attention to behaviorally relevant and task-related but currently unattended stimuli (Corbetta and Shulman, 2002). The importance of rTPJ in the bottom-up reorientation of attention, which is responsible for stimulus-driven changes in attentional focus, is well supported by the neurological condition of hemispatial neglect. Regardless of motor or sensory deficits, patients with spatial neglect show difficulties in spontaneously reorienting attention to stimulus information in the contralesional visual field. Importantly, damage to rTPJ is frequently linked to the occurrence of spatial neglect (Corbetta and Shulman, 2011), confirming the pivotal role of this structure in reorienting attention. In contrast to a fractionation view, recent attempts suggest common overarching mechanisms to integrate the variety of functions assigned to rTPJ. The nexus model (Carter and Huettel, 2013) tries to conciliate in a unified theory the role of rTPJ in mentalizing and attention, suggesting that TPJ is a “nexus” area where the overlap of lower-level functions (i. e., reorienting attention) leads to higher-order social-cognitive functions (i.e., mentalizing). The key idea of the nexus model is that TPJ would integrate information from different domains of cognition to construct a social context promoting decision-making. A further overarching view is known as the circuit-breaking theory by Corbetta et al. (2008); Corbetta and Shulman, (2002). This theory hypothesizes that the reorientation of attention is the result of the coordinated action of the ventral and the dorsal attention network. Broadly speaking, rTPJ plays a critical role in the detection of unexpected but task-relevant stimuli. As a node of the VAN, rTPJ would therefore act as a sort of circuit breaker for the dorsal attention network (DAN; Vossel et al., 2014). The activity of the latter, which is responsible for maintaining visuospatial information relevant to the current task-defined goals, would be interrupted by a signal of the VAN, with the consequence of reorienting the attention to a new salient stimulus. Noteworthy, this theory has been applied to social cognition as well, considering rTPJ as a relay to shift attention from an egocentric perspective to someone else’s perspective, and vice versa (Corbetta et al., 2008). An alternative explanation of the role of rTPJ in cognition is suggested by Geng and Vossel (2013), who refuse the core assumption of the circuit-breaking theory. The circuit-breaking theory posits that rTPJ triggers the reorienting of attention by interrupting the activity of the DAN. However, if it was true, this would entail that the activity of rTPJ occurs earlier than the activity of the brain regions of the DAN. On contrary, evidence from event-related potentials and transcranial magnetic stimulation (TMS) studies showed an opposite scenario: the frontal eye fields, which are part of the DAN, are activated earlier than rTPJ (i. e., the VAN) during attentional reorienting (Bardi et al., 2012; Thompson et al., 1996). Although Geng and Vossel (2013) agreed with the well-established role of TPJ in attentional control, they introduce an alternative explanation to reconcile the above-mentioned inconsistencies in an integrative view. Their theory originates from the functional meaning attributed to the P3 (also known as P300), an electrophysiological potential related to “contextual updating” (Polich, 2007). The P3 has been associated with the activity of several neural sources. The P3b, which is a particular subcomponent of P3, has been linked with TPJ and is considered to be a neurophysiological correlate of contextual updating (Polich, 2007; Verleger et al., 1994). In the contextual updating hypothesis, TPJ would update internal models of the environment/context to generate and drive adequate expectations and actions (Geng and Vossel, 2013). This theory is supported also by functional magnetic resonance imaging (fMRI) evidence showing the activation of TPJ in attentional tasks, in which the violation of expectations triggers the largest responses in this region (Vossel et al., 2006). Remarkably, the contextual updating hypothesis offers a model to interpret both low-level functions, for example reorienting attention, and high-level social functions, such as ToM, suggesting a common computational mechanism (i.e., contextual updating). From a social perspective, TPJ contributes to integrating the contextual representation of social situations needed to take the perspective of another person. Partially in line with the view of Geng and Vossel (2013), especially about the role of rTPJ as a hotspot involved in updating internal models, is the attention schema theory. This theory postulates a relationship between awareness and selective visual attention. Specifically, this framework states that awareness reflects an inner representation of selective attention, therefore a model of attention containing a set of constantly updated information about the current state of attention. This attention model, or “attention schema” - as the authors state - allows to monitor and control attention, predict future changes in attention level, and predict how attention can affect behavior and cognition (Graziano, 2020; Graziano and Kastner, 2011; Graziano and Webb, 2015). It is known that selective attention can be defined as exogenous (e.g., attention automatically captured by a sudden flash) and endogenous (e. g., attention voluntarily allocated to perform a task). When a prediction is not confirmed, that is when awareness and attention dissociate, a misalignment between the model of attention and the actual attention occurs, leading to an impairment of the endogenous control of attention. The violation of the prediction generates an error signal aimed at updating the model itself. As a recent fMRI study has demonstrated, rTPJ has a relevant role in this framework because it contributes to constructing and updating models of attention (Wilterson et al., 2021). The authors showed that when people are aware of a visual cue, they implicitly use the cue to generate awareness-dependent predictions aimed at driving endogenous attention to a target. Of interest, they argued that the brain area activated by the violation of predictions is rTPJ. In contrast to the circuit-breaking theory, rTPJ (and the VAN) would not serve exogenous attention. Instead, rTPJ would be implicated in creating or adjusting a model of attention. A surprising or unexpected event that captures attention away represents a violation of the attention model. In this circumstance, rTPJ has a rise in activity, as Wilterson et al. (2021) have shown. As the above-mentioned perspectives outline, clarifying the role of rTPJ in cognition has proven challenging because rTPJ is engaged across a broad range of processes. Taken together, both the perspectives attributing specific functions to rTPJ and the theories positing overarching mechanisms provide a fragmented view regarding the involvement of this region in cognition. This is going to translate this knowledge into practical solutions very difficult. For example, current neuropsychological measures and clinical treatments mostly reflect this theoretical fragmentation as they focus on assessing and treating, separately, cognitive or functional domains (e.g., language, attention, ToM), hence assuming a modular view of brain functioning (Fotopoulou, 2014). The modular assumption, despite being historically central in cognitive science (Fodor et al., 1983), may not be the unique rationale underlying the creation of neuropsychological tests and clinical interventions. Although it is well-known that after brain damage cognitive functions may show dissociations of deficits (supporting a modular view of the mind; Shallice, 1988), it is noteworthy to underline that most of the time clinical patients show associations of deficits (other than dissociations; Corbetta et al., 2018, 2015). The importance of association of deficits has also been acknowledged in the field of neuropsychological rehabilitation, as typically deficits do not occur in isolation (Wilson et al., 2017). Consequently, the adoption of an alternative view, different from the "modularist" tradition, may stimulate a rethink of the rationale of clinical treatments and provide different expectations on the treatment outcomes (Brown and Kuperberg, 2015). One particular framework that can support this paradigm shift is the predictive processing.

1.2. Predictive processing framework for brain disease

In the last decade, the concept of predictive processing has become popular in cognitive neuroscience because it outlines a unifying framework for understanding neural computation underlying perception, cognition, and action (Clark, 2016, 2013). Although several theories and algorithms have been developed within the "predictive processing" umbrella term (elegantly summarized here: Aitchison and Lengyel, 2017; Euler, 2018; Nave et al., 2020; van Elk, 2021), in this review we will refer to Andy Clark's general view (Clark, 2016, 2013). As mentioned, different frameworks have been proposed on the role of rTPJ in cognition, however, we decided to consider the deficits associated with rTPJ impairment under the lens of the predictive processing, as it provides an overarching framework that allows for interpretation of different aspects of cognition in terms of hierarchical processes of sensory inputs, prediction error, and model updating. Of importance, referring to Clark's framework does not mutually exclude the possibility of being in line with other models which consider the role of rTPJ as a structure responsible for the integration of information and updating (e. g., as suggested in the contextual updating hypothesis and the attention schema theory: Geng and Vossel, 2013; Wilterson et al., 2021), which fits well with the idea of a predictive brain (Clark, 2013; Friston, 2019, 2010). More in detail, Clark offers an integrative perspective of the abovementioned theories (Swanson, 2016) by postulating, through the predictive processing, that an internal model identifies causal regularities from a complex variety of sensory signals, extracting what is significant and salient to predict future events (Nave et al., 2020). The mismatch between a prediction and the actual sensory input generates a prediction error. The goal of the "predictive brain" is to try to generate more and more accurate predictions and therefore minimize the prediction error (Clark, 2013, 2016).

This prediction error propagates up the processing hierarchy to update expectations and provide better predictions that are transferred down the hierarchy (Friston, 2019; Keller and Mrsic-Flogel, 2018), although an alternative mechanism to hierarchical predictive processing has been proposed as well (Heeger, 2017). Along with estimating predictions about future events, the brain also estimates the reliability of these predictions and the reliability of the incoming sensory information or, in other terms, their precision (Doya et al., 2006; Yon and Frith, 2021). Precision refers to the level of confidence in predictions, in other words, the estimation of their occurrence likelihood. Concepts such as prior, posterior, and likelihood draw on the Bayesian brain hypothesis, a dominant - but also debated (Rahnev, 2019; Sanborn and Chater, 2016) - perspective to model inferential processes within the brain (Hohwy, 2017). Most of the time the causes of incoming sensory information are latent, simply unknown. Thus, the brain has to face an inverse problem, namely to estimate the causes of sensory information by generating inferences that result from both known information, so-called priors in Bayesian terms, and the incoming sensory information. The prediction resulting from priors and incoming sensory information is defined as the posterior probability that reflects the likelihood of the latent information being the cause of the sensory information. When either priors or sensory information are given more/less precision than they should, predictions about our reality may become less accurate or aberrant (Howes et al., 2020; Kube et al., 2020; Smith et al., 2021). From a theoretical point of view, any alteration of the "predictive brain" may cause a cascade of consequences leading to domain-general impairments. As suggested by Fotopoulou (2014), who interpreted anosognosia for hemiplegia as a deficit of predictive processing, several kinds of disruptions may impair the recursive dialectic between prior beliefs and current experience: (1) deficits to update predictions; (2) weak or absent prediction error signal; (3) firm adherence to predictions; (4) inability to optimize the precision of prediction errors; and, in general, (5) aberrant predictions. In line with Fotopoulou (2014), Kocagoncu et al. (2021) suggested that aberrant predictive mechanisms may explain impairments of neurological patients. In general, difficulties resulting from disruptions at any level of predictive processing may affect both the cognitive and the motor domains.

1.3. rTPJ as a key node in predictive processing

The hypothesis regarding the association between rTPJ and predictive processing is relatively recent but has been already supported by and discussed in several studies (Koster-Hale and Saxe, 2013; Park et al., 2021; Wilterson et al., 2021). For example, recently Park and colleagues (2021) investigated a potential causal link between rTPJ and the updating of predictions. In their study, the authors used fMRI to investigate the dynamic process underlying the maintenance or updating of social impressions of others, as a function of the violation of expectations (i.e., the prediction error). They showed that the recruitment of rTPJ reflected the integration of prediction error signals: specifically, the engagement of rTPJ was related to the processing of the social prediction error and updating of prior beliefs. Similarly, corroborating the centrality of rTPJ in predictive processing, Mengotti et al. (2017) applied online TMS over rTPJ. In line with their hypothesis, TMS interfered with the participants' capability to update prior beliefs. In addition, the TMS disruption was determined only in precise timing, namely when the TMS pulse was delivered 300 ms after a target stimulus, not at 50 ms. Of interest, the authors explained this result as further corroboration of the idea that rTPJ is one of the neural sources of the P3, an ERP component strictly associated with contextual updating (Geng and Vossel, 2013; Polich, 2007). An interesting view has been suggested by Siman-Tov et al. (2019) who conducted a coordinate-based meta-analysis of neuroimaging studies engaging both prediction generation and violation. Their results showed a set of cortical and subcortical brain regions, including the inferior and middle frontal gyri, anterior insula, premotor cortex, pre-supplementary motor area, striatum, thalamus, cerebellum, as well as rTPJ. This assembly of structures would be engaged in higher-level predictions that have been proposed to belong to a putative "prediction network" subserving both perception and action.

Interestingly, rTPJ is the only region within this prediction network that is also a crucial area of the VAN (Corbetta and Shulman, 2002). This concurrency has yielded the authors to suggest an intriguing role of rTPJ within the dynamics occurring between the VAN and the prediction network. Along with the well-established role of the VAN in the stimulus-driven changes in attentional focus, this network would be also activated in case of violation of expectations (Geng and Vossel, 2013; Wilterson et al., 2021), evidence that links the VAN to the predictive processing framework. Specifically, Siman-Tov et al. (2019) argued that the VAN might be part, together with other brain structures, of a larger network subserving domain-general high-order prediction. Pushing this idea further, the authors suggest that rTPJ, as a relevant hub of the VAN, has a key role in estimating the precision of priors and sensory evidence. This hypothesis is not new, since attention has been already interpreted through the lens of predictive processing (Feldman and Friston, 2010; Hohwy, 2013). However, this scenario provides a sharpened interpretation of neural correlates of mechanisms responsible for predictions. To sum up, a consistent number of studies outline a view in which rTPJ seems to be a fundamental cortical region subserving predictive processing and possibly related to a core set of areas identified as a “prediction network”. As we have seen before, rTPJ is certainly a crucial structure of the "predictive brain". However, advancing the hypothesis that direct damage or disconnection of this area is a necessary condition to disrupt the prediction network and, consequently, cause deficits ascribable to aberrant predictive processing would be, at the moment, merely speculative. Instead, neurological disorders affecting rTPJ may be a sufficient, even if not necessary, condition to cause difficulties related to predictive processes, leading to a plethora of impairments such as overconfidence in prior beliefs or underconfidence in sensory information, alongside a difficulty to detect errors, and a general inability to distill probabilistic structures in the world. Of course, predictive processing deficits can also emerge as a consequence of the impairment of other brain regions, in line with Siman-Tov et al. (2019) who suggest that rTPJ is only part of an extended network subserving high-level predictions. This study aims to provide a state-of-the-art overview of structural and functional aberrant outcomes following rTPJ “disconnection” from a putative prediction network, suggesting the involvement of rTPJ in a network potentially responsible for prediction.

2. Methods

The current systematic review is reported according to the recently updated guidelines of Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA, Page et al., 2021).

2.1 Search

A systematic literature search was performed on three databases: PubMed, PsycInfo, and Embase. The search terms included the conjunction of the following terms: ((“right” AND (“TPJ” OR “temporo parietal junction” OR “temporoparietal junction” OR “temporo-parietal junction” OR “temporo-parietal region”)) OR “rTPJ”) AND (“lesion*” OR “neurological” OR “Huntington” OR “mild cognitive impairment” OR “Parkinson disease” OR “multiple sclerosis” OR “stroke” OR “brain damage” OR “dementia” OR “Alzheimer” (all keywords) OR “stroke” OR “multiple sclerosis” OR “dementia” (all Mesh terms)). For details about the full search strategy for all databases please see the Supplementary Materials. Additional records relevant to the topic of the systematic search (e.g., articles that were cited by other articles) were also included. No restriction on publication date range was applied (last date of search: 23rd October 2020) and only studies with an English version and published in peer-reviewed journals were considered. Inclusion criteria were established a priori and were the following: (i) peer-reviewed papers; (ii) papers written in English; (iii) papers including adult (> 18 years) neurological patients; iv) papers studying the functional role of rTPJ (both due to a direct lesion of rTPJ or to its disconnection from other structures). Exclusion criteria were also established a priori and were the following: (i) papers without data analysis; (ii) reviews, meta-analysis, methodological papers; (iii) papers found incidentally by the algorithm, but not including rTPJ; (iv) papers involving animal models, children (< 18 years), healthy older adults only, non-neurological patients.

2.2 Risk of bias

To reduce the risk of bias, two independent reviewers (F.M. and S.L.) screened the articles and decided for appropriateness. Discrepancies in the evaluation of an article were resolved by a third author (R.P.). To assess the quality of the studies included, we applied the Newcastle-Ottawa Scale (NOS), a tool developed to evaluate non- randomized studies for systematic reviews (Wells et al., 2011); more specifically, we used a version adapted for cross-sectional studies (Patra et al., 2015) and we customized it according to the aims of our systematic search (please, see Table S1 and S2, in the Supplementary Materials, for details on NOS).

2.3 Data extraction

Data extraction was performed by F.M., R.P., and S.L. Each author was assigned a subset of papers to review. For each study, we highlighted the aim of the study, whether it focused on a cognitive domain and, if so, which specific domain was investigated, differentiating between *a priori* interests or incidental findings. This information was included to pinpoint potential selective biases of rTPJ studies towards certain cognitive functions. Data extracted from each subset of papers was subsequently and independently double-checked also by the authors that did not perform data extraction on that specific subset.

3. Results

The systematic literature search provided a total of 539 articles, as shown in the PRISMA flow diagram (Fig. 1). After discarding duplicates, a total of 199 records remained. These articles were screened based on title and abstract, according to the established criteria; 119 articles were considered not appropriate; thus, the full text of 81 articles was screened. A final set of 54 articles were judged suitable to be included in the review as the remaining 27 papers did not meet the inclusion criteria.

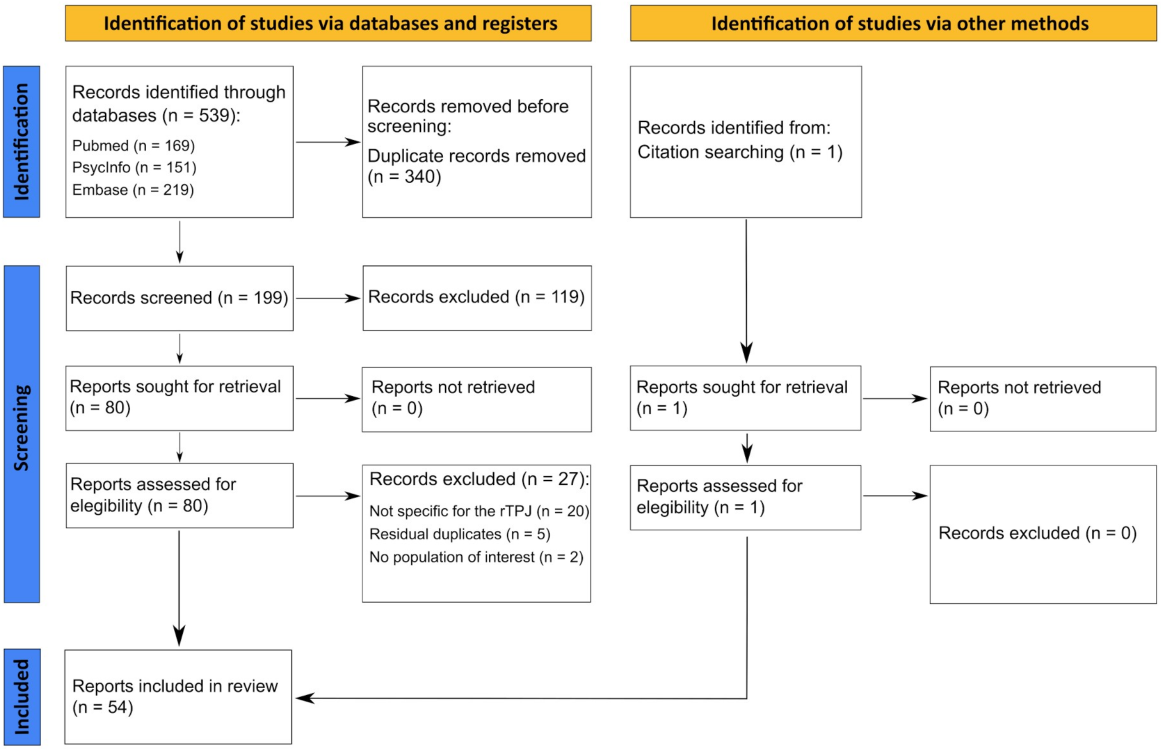


Figure 1. PRISMA flow chart outlining the process of article selection.

The 54 identified studies were clustered into the following 4 subgroups according to the pathology of interest: (i) “acquired brain injury” (which included right non-traumatic damage, bilateral non-traumatic damage, and traumatic brain injury); (ii) “neurodegenerative disease” (including Alzheimer, Mild Cognitive Impairment, Frontotemporal Dementia, Dementia with Lewy Bodies, PD); (iii) “white matter disease” (which included multiple sclerosis and hyperintense lesions); (iv) several pathologies that have been grouped as “other pathologies” as they referred to a minority of studies found (epilepsy, psychogenic nonepileptic seizures, migraine, and hemianopia). Importantly, we included both studies where the lesion involved focal damage to rTPJ and studies where rTPJ was not directly damaged but the lesion caused a disconnection between rTPJ and other brain areas/networks. In this latter case, the functional damage was interpreted by the authors as a consequence of the segregation of rTPJ from other areas/networks. A summary of the main results for each subgroup can be found in Table 1. Table S3 summarizes information about brain structures, in addition to rTPJ, involved in the selected studies, and whether rTPJ is directly damaged, disconnected, hypoperfused, or other forms of alteration. In each of the following sections, we present a synthesis of the main results for each subgroup.

3.1 Acquired brain injury

Acquired brain injury is an umbrella term encompassing a wide spectrum of brain lesions that are not hereditary, congenital, or degenerative, and are characterized by a traumatic/non-traumatic etiology. Acquired brain injury can result in impairments in cognition, motor function, sensory processing, and emotional disturbances. Studies on patients with acquired brain injury can help investigate how rTPJ damage affects cognitive functioning. In the present review, 29 studies have been included within this subgroup: 22 studies regarding right nontraumatic brain lesions, 6 studies with bilateral non-traumatic brain damage, and 1 study with patients with traumatic brain injury (TBI). Sixteen out of 29 studies specifically mentioned patients with hemispatial neglect, and their results have been summarized at the end of this section. Several authors suggest that cognitive and motor impairments observed in patients could result both from direct damage of rTPJ and from a disconnection of this brain structure from other areas or networks. This latter scenario implies that rTPJ can be directly damaged or not. For example, the disconnection of rTPJ from the frontoparietal network could be responsible for a detrimental effect on orienting attention to relevant events (Pedrazzini and Ptak, 2019), anosognosia for hemiplegia (Monai et al., 2020), and a selective impairment at the expense of cognitive flexibility (i.e., executive functions), specifically in set-shifting abilities (Mandonnet et al., 2017). Interestingly, Wawrzyniak et al. (2018) linked the patients' deficits to a breakdown of networks of the right hemisphere.

Through the lesion network-symptom-mapping approach, the authors investigated the phenomenon of the rubber hand illusion in patients with mild to moderate stroke symptoms and demonstrated the engagement of rTPJ in eliciting the illusory body-ownership feeling. A similar result was found by Martinaud et al. (2017) who showed that lesions associated with disturbing sensations of limb ownership included rTPJ, supramarginal gyrus, and middle frontal gyrus. Thus, rTPJ, together with other brain regions, seems to support the integration of body sensations to achieve a global body representation (Boccia et al., 2020). Studies included in the next subsection found an association between rTPJ and attention impairment, without advocating a disconnection of this area with other circuits but interpreting deficits as the consequence of rTPJ damage itself. For example, Shomstein et al. (2010) studied the neural bases of top-down (goal-directed) and bottom-up (stimulus-driven) attentional orienting. Starting from previous evidence showing the involvement of the superior parietal lobule for top-down attentional orienting and rTPJ for bottom-up attentional orienting, they selected a group of patients with lesions in these regions of interest. A lesion overlap analysis showed a double dissociation: (1) impairment of top-down attentional orienting, but normal bottom-up orienting, was related to superior parietal lobule lesions; (2) impairment of bottom-up attentional orienting, but normal top-down orienting, was related to TPJ lesions. Together with attentional orienting, rTPJ seems to play a role also in temporal attention. Agosta et al. (2017) investigated the capability to perceive the sequential order of two events, asking participants to perform temporal judgment tasks. To this aim, they conducted two experiments on a group of patients with left and right parietal lesions. Results of the first experiment showed reduced visual temporal processing in patients with right parietal damage, while patients with lesions in the opposite hemisphere performed normally. The second experiment confirmed the role of the right parietal brain regions in temporal judgments. In particular, repetitive TMS over rTPJ - but not the left TPJ - bilaterally disrupted the ability of participants to visually discriminate stimuli across time. Along with attention, rTPJ has been found to be implicated in a variety of other functions. For instance, Starkstein et al. (1992) investigated the neural correlates of anosognosia, showing that patients with mild or severe anosognosia more often had temporoparietal lesions (among several other lesions), as compared to patients with no or moderate anosognosia. As for the relation of rTPJ with mental state attribution, in Leigh et al. (2013) patients with right acute stroke were recruited to investigate whether deficits in affective empathy were related to lesion site alone, or also affected by other aspects, specifically lesion volume, age of patients, neglect, and prosody comprehension. What the authors found, confirming previous evidence, is that a unilateral lesion of rTPJ is not sufficient to impair affective empathy. Similarly, Cohen-Zimerman et al. (2021) employed voxel-based lesion-symptom mapping to examine the relationship between focal brain lesions and mental state attribution. They tested a large sample of patients selected from the Vietnam Head Injury Study. In line with Leigh et al. (2013), Cohen-Zimerman et al. (2021) did not reveal the involvement of rTPJ in mental state attribution. Thus, both these studies suggest that rTPJ could not be necessary for mental state attribution. Finally, rTPJ was also associated with motor functions. Indeed, several studies revealed the role of this brain region in motor planning and sensorimotor integration. Singh and Knight (1993) aimed to identify a relation between movement-related potentials, measured in a self-paced button-press task, and unilateral lesions in the posterior association cortex. Results from this study revealed an involvement of the posterior association cortex damage in reducing the movement-related potential amplitudes. This result established the role of the superior parietal regions (and rTPJ) in movement preparation. Kaski et al. (2016) studied the neural underpinnings of vestibular-spatial perception asking patients with right damage to perform a series of vestibular reorientation tasks in the dark. This study showed that the brain encodes self-motion and spatial perception separately and rTPJ seems selectively crucial in spatial perception. The authors suggested that this area may work as a mental temporal integrator somehow, in which the subjective estimation of motion velocity over time is integrated with the updating of traveled distance perception in the dark (i.e., under vestibular guidance). As previously mentioned, in the section related to acquired brain injury, 16 papers explicitly considered hemispatial neglect. Lesion analysis confirmed that rTPJ is associated with the occurrence of spatial neglect (Demeurisse et al., 1997; Ptak and Schnider, 2011). Crucially, rTPJ contributed to predicting the patients' performance in a variety of visuospatial tasks (Toba et al., 2020). For example, Kaufman et al. (2009) found that line bisection deficits in patients with neglect were associated with damage to the anterior part of rTPJ, among other temporal regions. In addition to gray matter damage, Golay et al. (2008) highlighted that patients who showed large bisection bias and small cancellation errors also had white matter damage located in the proximity of rTPJ. Similarly, Thiebaut de Schotten et al. (2014) investigated the role of white matter disconnection in chronic spatial neglect measured with line bisection and letter cancellation. Results of their study revealed that white matter lesions in rTPJ, among other frontoparietal areas, predicted whether a patient showed neglect symptoms and had a worse bisection performance. Pedrazzini et al. (2017) enrolled neglect patients with right-hemispheric lesions to examine which sites were more strongly associated with visuospatial processing of single objects or space. They observed that damage to rTPJ was a strong predictor of space-based variables only, suggesting that space and object-based processing have distinct neural bases and that space-based processing is grounded in rTPJ. Additional typical manifestations of neglect encompass the phenomenon of visual extinction, which is the inability of brain-damaged patients to detect a contralesional target in the presence of a competing ipsilesional stimulus. In a study by Karnath et al. (2003), the authors classified patients into three groups according to their symptoms: pure extinction (no neglect), extinction plus neglect, and pure neglect (no extinction). Results showed that patients with extinction and neglect had lesions comparable to that of pure extinction patients, namely at the level of rTPJ, covering also other temporoparietal areas. Instead, extinction plus neglect patients reported greater damage of rTPJ compared to those with extinction only. Also, Ticini et al. (2010) focused on visual extinction, conducting normalized perfusion-weighted MRI analysis. Findings indicated that patients with visual extinction showed cortical malperfusion around rTPJ. Altogether, these findings lend support to the hypothesis that rTPJ is essential for the bottom-up detection of stimuli. Together with visuospatial processing, it has been found that rTPJ is implicated in a variety of other functions. Among these, Chechlacz et al.

(2014) showed that areas typically associated with neglect may support visuospatial memory performance. In a recent study, Dressing et al. (2020) investigated the relationship between neglect and apraxia (an impairment in tool use or imitation of gestures, naturalistic actions, and meaningless postures) in patients with right hemisphere damage. Interestingly, voxel-based lesion-symptom mapping revealed a negative correlation between apraxia and rTPJ integrity. Others investigated the anatomical substrates of neglect dyslexia showing the pivotal contribution of rTPJ in reading (Lee et al., 2009). Finally, Rousseaux et al. (2015) examined the neural correlates of behavioral manifestations of neglect in daily activities. Findings revealed that difficulties in daily living mainly resulted from subcortical white matter lesions underlying rTPJ, among other temporoparietal areas. A separate mention should be made for all those studies that have investigated how rTPJ disconnection from other areas or networks can yield different consequences. Hattori et al. (2018), for example, showed that in patients with rTPJ lesions the VAN was damaged. Lesions included the white matter connecting rTPJ and surrounding areas with frontal regions. Moreover, in those patients with infarction in the territory of the posterior cerebral artery, white matter tracts connecting the thalamus to rTPJ and the surrounding areas were damaged. The authors argued that different lesion patterns may underlie different neglect phenotypes. Specifically, damage to the VAN impairs stimulus-driven, bottom-up attentional reorienting from current focus to new and unexpected focus by modulating the DAN. Instead, damage in the thalamus and related white matter tracts may secondarily impair the spatial attention-related cortices due to interrupted somatosensory and feedback information from the thalamus. In another study, Committeri et al. (2015) tested patients for representational neglect, visual extrapersonal/perceptual neglect, and personal neglect. Results showed that only rTPJ emerged as significantly more involved in the genesis of representational neglect for places. In particular, the posterior rTPJ appeared uniquely connected with the DMN. Finally, Pedrazzini and Ptak (2020) investigated the neural bases of spatial awareness, which is particularly impaired in neglect. Results showed that damage at the supramarginal part of rTPJ, together with frontoparietal connections, was the best predictor of neglect. To summarize, regarding the functional alteration investigated by each study, acquired rTPJ damage was associated with impairment of attention, awareness, executive functions, visuospatial memory, spatial perception, movement preparation, and difficulties in activities of daily living (see Table 1). These deficits have been interpreted in two ways: locally or globally. The local explanation considers the functional impairment as a direct consequence of the rTPJ lesion. Instead, the global explanation outlines a more complex scenario where the rTPJ damage, together with the alteration to other brain structures, is the cause of the functional impairment.

3.2 Neurodegenerative diseases

Neurodegenerative diseases are a group of disorders characterized by the progressive degeneration of nerve cells, in the central or the peripheral neural system (Palop et al., 2006). Fifteen studies were included in this subgroup, 13 studies comprise patients with various types of dementia or cognitive impairment, including Alzheimer’s Disease (AD), frontotemporal dementia (FTD), mild cognitive impairment (MCI), functional neurological disorder, corticobasal syndrome, dementia with Lewy bodies (DLB). Two out of 15 were conducted on patients with PD. Among the studies that investigated the relationship between rTPJ and attention in AD patients, the one by Sorg et al. (2012) found that direction and degree of spatial attention bias correlated with TPJ activity. Furthermore, Yamashita et al. (2019) showed that the connectivity between rTPJ and posterior parietal cortex was diminished in poor performers during a task on orientation for time. In another study, Luks et al. (2010) recruited a heterogeneous group of patients with several neurodegenerative disorders. Results showed that the atrophy in the TPJ-ventrolateral prefrontal cortex (VLPFC) network was associated with slower attentional control in accurate trials. This network (i.e., TPJ-VLPFC) is thought to be responsible for reorienting attention towards salient and infrequent stimuli, in a bottom-up fashion. In the absence of an efficient top-down attentional control system mediated by the DLPFC-ACC network, accurate and speedy processing of incongruent Flanker task stimuli may be accomplished by the bottom-up reorienting and inhibitory mechanisms of the TPJ-VLPFC network. A limited number of studies also focused on aspects related to selfawareness, such as Pickut et al. (2013) who found increased gray matter density in a variety of regions, including bilateral TPJ, in a group of PD patients who followed an 8-week mindfulness-based intervention compared to standard training. Concerning self-awareness, Zamboni et al. (2010) found that in patients with various forms of FTD, the degree of anosognosia for behavioral impairment (which is a condition in which patients are often partially, if not completely, unaware of the behavioral deficits) correlated with gray matter atrophy in the posterior regions of the brain, near rTPJ region, suggesting a role for rTPJ in anosognosia and general self-awareness. Also, Baez et al. (2019) tested patients with behavioral variant FTD (bvFTD) and patients with bipolar disorder, which present overlapping symptoms with the neurodegenerative disorder in the domain of cognitive and social functioning. BvFTD patients, who also showed greater executive functions and theory of mind deficits than psychiatric patients, had atrophy in several regions, including rTPJ. The authors suggested that rTPJ is intended as part of the extended cortical-limbic networks, which play a role in social cognition. As for memory deficits, Kang et al. (2019) investigated the effects of AD and DLB on cognition and brain atrophy, and they found that AD was associated with prominent memory deficits and brain atrophy in the medial temporal lobe and temporoparietal association cortices, while DLB was characterized by visuospatial, attention, and executive dysfunction. The majority of studies found an impairment of rTPJ while investigating the structural and functional integrity of the brain, without a priori focus on a specific cognitive domain. Qian et al. (2015) investigated the functional and structural substrates of attention by testing the integrity of the VAN and DAN networks. Their data showed decreased functional connectivity in the orbital ventral frontal cortex and TPJ region in AD patients, whereas functional connectivity was preserved in amnestic MCI compared with controls. In addition, gray matter density in the right ventral frontal cortex was correlated with functional deterioration in rTPJ and ventral frontal cortex in AD. Recently, De Marco et al. (2019) used structural MRI data of AD, MCI patients, and controls to disentangle DMN patterns that are associated with normal aging or neurodegeneration. They found a significant positive association between hippocampal volumes and the DMN connectivity in rTPJ. The authors suggested that rTPJ may subserve ToM processes and social cognition through its connection to the DMN. They also found an association between rTPJ and verbal episodic memory. Diez et al. (2019) observed that altered connectivity in rTPJ (as part of the multimodal integration network) and insular regions correlated with neurological symptoms, including altered interoception and self/emotional awareness. These results suggested that the functional alteration in insula and rTPJ play a crucial role in promoting altered awareness in functional neurological disorders. In addition, the study of Zou et al. (2014) revealed a decrease in cerebral blood flow values in AD patients compared to controls in bilateral frontal regions, the temporal lobe, bilateral TPJ, as well as parietal and hippocampal regions. In a longitudinal investigation, a significant correlation between the worsening of depression and the lowering of the cortical thickness in rTPJ over time (after approximately 18 months) was observed in non-demented PD patients (Hanganu et al., 2017). Three studies investigated the role of apathy in patients with neurodegenerative disorders, finding structural or functional implications of rTPJ. Along this line, Blanc et al. (2015) investigated cortical thickness in a variety of patients with neurodegeneration. In the patients with DLB, the cortical thinning was found predominantly in rTPJ, insula, and other cortices, as compared to controls. In another study, Rohrer et al. (2012) examined patients with primary progressive aphasia with AD or the presence of cerebrospinal fluids compatible with AD, showing a pattern of cortical thinning in many regions when patients were compared to controls. In those patients with more severe diseases, they found increased involvement of the left anterior temporal and frontal cortices, posterior cingulate, medial temporal lobe, and rTPJ. Another study had a similar aim, namely to investigate the cortical changes in association with apathy (Eslinger et al., 2012). They showed that apathy was significantly correlated with atrophic changes in several regions, including rTPJ in patients with bvFTD, compared to other forms of dementia. Behaviorally, multiple measures of executive function and social cognition were impaired in the bvFTD sample. To sum up, this paragraph shows how the involvement of rTPJ in neurodegenerative disorders is associated with a variety of impairments. Functionally, deficits associated with rTPJ consisted of impairment of attention, social cognition, executive functions, awareness (anosognosia), memory, and language. In addition, several studies showed the relationship between rTPJ, mood, and apathy (see Table 1).

3.3 White matter disease

In this paragraph, we discuss studies on a group of heterogeneous disorders that engage degeneration of the white matter of the brain, which could disrupt cognitive functioning or result in other symptoms. Studies on patients with white matter disruptions can help investigate whether and how structural disconnections also impact rTPJ functioning. Five studies have been included in this subgroup, 4 studies with patients with multiple sclerosis (MS), and 1 study on patients with white matter hyperintense lesions. In a first study, Carotenuto et al. (2018) tested MS patients, who showed a correlation between rTPJ and deficits in pragmatic and communicative abilities (tested with the Assessment of Pragmatic Abilities and Cognitive Substrates - APACS - test, Arcara and Bambini, 2016). Here, a direct correlation between the APACS score and a cluster within the paracingulate cortex was found when evaluating the right Geschwind's area (encompassing rTPJ), thus showing a relation between this area and pragmatics abilities. Huang et al. (2019) tested patients with relapsing-remitting MS that showed decreased dynamic functional connectivity within both the DAN and VAN, of which rTPJ is part, but increased connectivity between the two networks; they also found that the connectivity between rTPJ and right ventral frontal cortex was negatively correlated with the total white matter lesion loads, suggesting that the detected effects of transient connectivity pattern are relevant to the lesions damage (Huang et al., 2019). Kim et al. (2019) tested resting-state connectivity in MS patients using magnetoencephalography. They reported an increase in alpha-band neural power in rTPJ, particularly in a subgroup of patients who experienced mixed neuropathic pain; the authors suggested that this effect may reflect an overactivity of rTPJ for overflowing sensory information. In another study, altered connectivity of rTPJ with the DMN was found in a subgroup of MS patients with mixed-neuropathic pain (Bosma et al., 2018). This result was interpreted as the index of an altered engagement of the ascending nociceptive and descending modulation pathway. Peng et al. (2016) quantified with voxel-based morphometry the gray matter density and correlated it with the white matter hyperintense load, this latter visually estimated by two expert neurologists. The authors found that patients with white matter hyperintense lesions had a greater increase in gray matter density in rTPJ. They also found a negative correlation between rTPJ and white matter lesion load (Peng et al., 2016).

3.4 Other pathologies

Five papers studied the involvement of rTPJ in pathologies not included in the previous sections (see Table 1). The first two studies focused on epileptic patients. Beauchamp et al. (2012) investigated the perception of phosphenes following electrical stimulation of the visual cortex, in patients with subdural electrodes for surgical treatment of epilepsy. Electrodes in the visual cortex provided the stimulation, whereas the non-stimulating electrodes implanted in other sites of the brain were used for recording. Results showed that a much greater response of gamma oscillations (60–150 Hz) was observed in TPJ and close areas during electrical stimulation of the visual cortex, particularly when phosphenes were perceived. Jiang et al. (2018) aimed at assessing the functional and causal connectivity patterns of the attention networks and DMN in patients with refractory epilepsy, using fMRI. All patients showed a decreased activation within the VAN in the bilateral TPJ and prefrontal cortex and the interaction between the attention networks (VAN, DAN) and DMN was altered, as compared to controls. One of the remaining four studies considered psychogenic nonepileptic seizures. Peterson et al. (2018) stimulated a small group of patients with psychogenic non-epileptic seizures with repetitive TMS over rTPJ, finding a decrease in weekly seizure rates after treatment. Improvement was also associated with lower measures of dissociation (particularly, conversion disorder). One study focused on the administration of transcutaneous auricular vagus nerve stimulation (taVNS) in migraine patients (Zhang et al., 2019). Results showed that functional connectivity in rTPJ and other sites (hippocampal, temporal, parietal, and pontine regions) significantly increased following real compared to sham stimulation. Other significant results included increases in resting-state connectivity between rTPJ and locus coeruleus, and between rTPJ and left secondary somatosensory cortex. Such increases were negatively associated with the frequency of migraine attacks during the preceding month. Finally, Lu et al. (2018) tested chronic hemianopic patients during visual rehabilitation training. Besides improvements in their contrast sensitivity (no behavioral attention tests were administered), functional connectivity results show increased activation in rTPJ after training and strengthening of connectivity between cingulate and insular regions to rTPJ (as part of the attention network).

Table 1. A summary of the studies included in the systematic review.

Authors	**Year**	**Damage**	**Pathology**	**Participants**	**Neuroimaging method**	**Attention**	**Awareness and social cognition**	**Language**	**Executive functions**	**Memory**	**Other domains**	**Behavioral measurements**	**Neuroimaging measurements**	**Findings (related to the rTPJ)**
Agosta et al.,	2017	Bilateral damage	Acquired brain injury	Experiment 1: 18 RD, 5 LD, 18 HC; Experiment 2: 10 HC	MRI and NIBS	AIM temporal attention	-	-	-	-	-	Computerized task: simultaneity judgment task	Repetitive transcranial magnetic stimulation	Results from the first experiment showed reduced visual temporal processing in patients with the right parietal damage, instead of patients with damage in the opposite hemisphere that performed normally. The second experiment confirmed the role of the right parietal brain regions in temporal judgments. In particular, repetitive TMS over the rTPJ - but not homologous one - bilaterally disrupted the ability of participants to visually discriminate stimuli across time
Boccia et al.,	2020	Right damage	Acquired brain injury	26 RD, 39 HC	MRI	-	AIM body structural representation	-	-	-	-	Computerized battery of tasks developed by the group to test body representations	Voxel-based lesion-symptom mapping	The rTPJ, together with other areas, may be crucial in the integration of body sensations to achieve a global body representation
Cohen-Zimerman et al.,	2020	Traumatic brain injury	Acquired brain injury	16 rTPJ, 7 lTPJ, 30 rDLPFC, 28 lDLPFC, 34 with no lesion in the PFC or the TPJ bilaterally, 30 HC	MRI	-	AIM theory of mind	-	-	-	-	The strange stories test	Voxel-based lesion-symptom mapping	The rTPJ was not involved in mental state attribution
Kaski et al.,	2016	Right damage	Acquired brain injury	18 RD, 2 avestibular patients, 14 HC	MRI	-	-	-	-	-	AIM vestibular-spatial perception, travelled distance, motion duration	Subjective angular position, velocity, and motion duration during whole-body angular rotations in the dark	Voxel-based lesion–symptom mapping	Patients with rTPJ damage showed impaired spatial orientation performance. This group of patients underestimated both the travelled distance and the motion duration during a task performed in the dark
Leigh et al.,	2013	Right damage	Acquired brain injury	28 RD, 24 HC	MRI	-	AIM affective empathy	-	-	-	-	Affective and cognitive empathy measured with several tasks: the Affective Empathy Task (stories), the Interpersonal Reactivity Index, the Aprosodia Battery	Diffusion weighted image	A unilateral lesion of the rTPJ is not sufficient to impair affective empathy
Mandonnet et al.,	2017	Right damage	Acquired brain injury	Single case	fMRI	-	-	-	AIM set-shifting abilities	-	-	Several tasks to explore: language functions, praxis, non-verbal semantic association, calculus, memory, attention, spatial awareness, executive functions	Resting state fronto-parieto-temporal networks extracted by means of ICA	The surgical disconnection of the rTPJ from the fronto-temporo-parietal network led to a selective impairment of cognitive flexibility, in particular the deterioration in set-shifting abilities
Martinaud et al.,	2017	Right damage	Acquired brain injury	31 RD	MRI	-	AIM body perception	-	-	-	-	A modified version of the Cutting questionnaire	Voxel-based lesion–symptom mapping	Damage associated with disturbed sensation of limb ownership included, among several areas, the rTPJ

												(Cutting, 1978)		
Monai et al.,	2020	Bilateral damage	Acquired brain injury	35 hemiplegia, 28 anosognosia for hemiplegia	MRI	-	AIM anosognosia for hemiplegia	-	-	-	-	-	Disconnectome maps	The disconnection between the inferior frontal gyrus with the TPJ may be critical in anosognosia for hemiplegia because this latter seems to integrate multimodal signals (body and visuospatial signals) and switch from inner to external perspectives
Pedrazzini and Ptak,	2019	Right damage	Acquired brain injury	10 rTPJ, 9 lPFC/insula, 10 subcortical damage, 12 HC	MRI	AIM orienting attention	-	-	-	-	-	Computerized task: spatial cueing task	Voxel-based lesion-symptom mapping	Patients with TPJ damage, but not the group with lateral prefrontal cortex/insula damage, exhibited exaggerated attention to behaviorally relevant cues
Shomstein et al.,	2010	Right damage	Acquired brain injury	9 RD, 9 HC	MRI	AIM attentional orienting	-	-	-	-	-	To examine the integrity of top-down attentional orienting: a variant of the Sperling and Reeves (1980) task. To examine the bottom-up attentional orienting abilities: a variant of Folk et al.'s (2002) contingent paradigm	Lesion overlap analysis	rTPJ damage was associated to impairment of bottom-up attentional orienting (but normal top-down orienting)
Singh and Knight,	1993	Bilateral damage	Acquired brain injury	7 TPJ, 5 lateral parietal lesion, 5 lateral parietal and TPJ damage, 14 HC	EEG	-	-	-	-	-	AIM self-initiated movements	Self-paced switch closures	Movement-related potentials	Role of the superior parietal regions (and the rTPJ) for movement preparation
Starkstein et al.,	1992	Bilateral damage	Acquired brain injury	80 BD (stroke)	CT	-	AIM anosognosia	-	-	-	-	Anosognosia questionnaire	Lesion volume	Patients with mild or severe anosognosia typically had temporoparietal damage compared to patients with no or moderate anosognosia
Wawrzyniak et al.,	2018	Bilateral damage	Acquired brain injury	10 patients not experiencing the RHI (RHI-), 37 experiencing the RHI (RHI+), 40 HC	fMRI	-	AIM illusory body-ownership feeling	-	-	-	-	RHI paradigm	Lesion network-symptom-mapping	Differences between who suffers from RHI failure and not were evident in the rTPJ, demonstrating the involvement of the rTPJ in eliciting the illusory body-ownership feeling
Chechlacz et al.,	2014	Bilateral damage	Acquired brain injury (neglect)	57 BD, 100 HC	MRI	AIM spatial attention	-	-	-	AIM working memory	-	Corsi Block Task, visual extinction, and Apple cancellation task	Voxel-based morphometry; tract-wise lesion deficit analysis	Areas typically associated with neglect (e.g., the rTPJ) may support visuospatial memory performance
Committeri et al.,	2015	Right damage	Acquired brain injury (neglect)	40 RD	fMRI	AIM visuospatial attention	-	-	-	-	-	Representational neglect: Squares test (familiar places); familiar objects (O'clock test); Visual extrapersonal neglect (Letter Cancellation Test, Line Cancellation Test, Wundt-Jastrow Area Illusion Test, Sentence Reading); personal neglect (semi-structured	Voxel-based lesion-symptom mapping	The TPJ emerged as significantly more involved in the genesis of representational neglect for places

												scale)		
Demeurisse et al.,	1997	Right damage	Acquired brain injury (neglect)	28 RD, 14 HC (only for behavioral normative data)	SPECT	AIM spatial attention	-	-	-	-	-	Letter cancellation task, Copy of Rey's figure, Albert's cancellation task, Spontaneous drawing of a wheel, Zazzo's cancellation task, Line bisection task, Reading of composite words	Cerebral blood flow	Neglect patients could be distinguished from patients without neglect by the presence of a remote decrease in cortical cerebral blood flow in the right temporoparietal region
Dressing et al.,	2020	Right damage	Acquired brain injury (neglect)	138 RD patients, 29 HC (only for behavioral normative data)	MRI	AIM spatial attention	-	-	-	-	AIM apraxia	Apraxia: imitation of meaningless gestures, pantomime of tool use; neglect: observation for signs of neglect, wiggle test, bells cancellation, line bisection, Ota test	Voxel-based lesion-symptom mapping	Imitation of meaningless postures was related to various lesion clusters, one of those is located in the rTPJ
Golay et al.,	2008	Right damage	Acquired brain injury (neglect)	50 RD (28 with and 22 without neglect)	MRI	AIM spatial attention	-	-	-	-	-	Line cancellation, bisection and copying of drawings tests	Voxel-based lesion-symptom mapping	The region differentiating best between neglect and control patients reached far into the white matter beneath the TPJ and STG
Hattori et al.,	2018	Right damage	Acquired brain injury (neglect)	59 RD (34 with and 25 without neglect), 5 HC	MRI	AIM spatial attention	-	-	-	-	-	Line cancellation, line bisection, cube copying test, clock drawing test	Non-parametric voxel based analysis	In patients with rTPJ lesions the VAN was damaged
Karnath et al.,	2003	Right damage	Acquired brain injury (neglect)	27 RD, 7 HC	MRI	AIM visual extinction and spatial attention	-	-	-	-	-	-	Lesion overlap analysis	Lesion overlap analysis showed a relation between visual extinction and lesions at the rTPJ
Kaufman et al.,	2009	Right damage	Acquired brain injury (neglect)	23 RD	MRI	AIM spatial attention	-	-	-	-	-	Line bisection	Multiperturbation analysis approach for lesion-symptom mapping	Among several regions, the rTPJ predicted line bisection performance
Lee et al.,	2009	Right damage	Acquired brain injury (neglect)	30 patients with neglect dyslexia, 50 patients without neglect dyslexia	MRI	AIM neglect dyslexia	-	-	-	-	-	Neglect: line bisection, Character Line Bisection Task, modified Albert's line cancellation test, Star Cancellation task, figure copying (modified Ogden Scene test, Two daisy figure); Neglect Dyslexia: single or compound words that varied in length from two to six syllables	Lesion location analysis	Lesions that were often associated with neglect were located in the region of the rTPJ, but patients with neglect dyslexia had on average larger lesions compared to the patients with neglect only
Pedrazzini and Ptak,	2020	Right damage	Acquired brain injury (neglect)	134 RD	MRI	AIM spatial attention	-	-	-	-	-	Bells cancellation test, letter cancellation, line bisection, and reading compound words	Voxel-based lesion–symptom mapping	One of the best predictors of impaired spatial exploration was a selective damage to the supramarginal part of the rTPJ
Pedrazzin	2017	Right	Acquired brain injury	101 RD	MRI	AIM spatial	-	-	-	-	-	Bells cancellation, letter cancellation, line	Lesion site	The TPJ was the only strong predictor of space-based processing

i et al.,		damage	(neglect)			attention						bisection, reading of compound words		
Ptak and Schnider,	2011	Right damage	Acquired brain injury (neglect)	29 RD (20 with and 9 without neglect), 10 HC	MRI	AIM spatial attention	-	-	-	-	-	Computerized spatial cueing task	Voxel-based lesion-symptom mapping	The TPJ is critical for the occurrence of spatial neglect but does not explain participants' performance in the cueing task
Rousseaux et al.,	2015	Right damage	Acquired brain injury (neglect)	45 RD	MRI	AIM spatial attention	AIM anosognosia	-	-	-	AIM activities of daily living	Peripersonal neglect: line bisection and target cancellation; personal neglect: reaching, test of the subjective straight-ahead; anosognosia for behavioral difficulties: discrepancy between clinician's and patient's ratings on Catherine Bergego Scale; anosognosia for hemiplegia or visual perception deficits: Bisiach's test for anosognosia	Voxel-based lesion-symptom mapping	Difficulties in daily living resulted from subcortical white matter lesions underlying the rTPJ, among other temporoparietal areas
Thiebaut de Schotten et al.,	2014	Right damage	Acquired brain injury (neglect)	58 RD (38 with and 20 without neglect)	MRI	AIM spatial attention	-	-	-	-	-	Line bisection, letter cancellation	Voxel-based lesion–symptom mapping, standard diffusion tensor imaging for single cases	Significant involvement of the TPJ, among other areas, in line bisection performance, but no involvement of the TPJ for letter cancellation performance
Ticini et al.,	2010	Right damage	Acquired brain injury (neglect)	13 patients (8 visual extinction+neglect, 5 neglect)	MRI	AIM visual extinction and spatial attention	-	-	-	-	-	-	Normalized perfusion-weighted MRI	Patients with visual extinction showed cortical malperfusion around the rTPJ
Toba et al.,	2017	Right damage	Acquired brain injury (neglect)	25 RD	MRI	AIM spatial attention	-	-	-	-	-	Line bisection, bells cancellation, letter cancellation	Multiperturbation Shapley value Analysis	The TPJ had positive interactions with a number of other areas and this contributed positively to the performance in the three visuospatial attention tests
Baez et al.,	2019	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	16 bvFTD, 13 bipolar disorder, and 22 HC	MRI	-	INC theory of mind	-	AIM executive functions	-	-	Reading the Mind in the Eyes test	Voxel-based morphometry	bvFTD and bipolar patients showed an altered structural activity in diverse regions; in bvFTD patients atrophy (i.e., reduction of gray matter) of several regions (including the rTPJ) was associated with theory of mind impairments
Blanc et al.,	2015	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	28 proDLB, 27 proAD, 31 DLB, 54 AD, 33 HC	MRI	-	-	-	-	-	-	-	Cortical thickness	In the patients with DLB the cortical thinning was found predominantly in the rTPJ, insula and other cortices, when compared to HC
De Marco et al.,	2019	Neurodegeneration	Neurodegenerative disease (dementia or cognitive	86 MCI, 35 AD, 191 HC	fMRI	-	-	-	-	INC verbal episodic memory	-	Prose Memory Test	Functional connectivity	A positive association was found between hippocampal volumes and default mode network connectivity in the rTPJ; an association between the rTPJ and verbal episodic memory was found

			impairment)											
Diez et al.,	2019	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	30 patients with motor FND, 30 HC	Resting-state fMRI	-	INC awareness	-	-	-	-	Screening for Somatoform Symptoms Conversion Disorder subscale and patients Health questionnaire	Graph-theory stepwise functional connectivity	FND patients had increased connectivity from motor regions to the bilateral posterior insula, TPJ, middle cingulate cortex and putamen. Interestingly, symptoms severity correlated with enhanced SFC from the left anterior insula to the right anterior insula and TPJ, supplementary motor area and sensorimotor areas
Eslinger et al.,	2012	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	26 FTD (12 bvFTD, 7 progressive non-fluent aphasia, 7 semantic dementia), 16 HC	MRI	-	-	-	-	-	AIM apathy	Apathy Evaluation Scale	Voxel-based morphometry	Apathy was significantly correlated with atrophic changes in several regions, including the rTPJ in the subgroup of bvFTD
Kang et al.,	2019	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	26 ADCI (18 MCI and 8 dementia), 28 LBCI (13 MCI and 15 dementia), and 54 mixed ADCI and LBCI (17 MCI and 37 dementia), 38 HC	MRI	AIM attention, visuospatial abilities	-	AIM language	AIM executive functions	AIM Verbal learning, memory	-	Seoul neuropsychological screening battery	Cortical thickness	ADCI and LBCI groups were independently associated with cortical thinning in the association cortices, including the bilateral TPJ, medial and lateral parietal, and DLPFCs cortices. The mixed group, namely the group that had concomitant ADCI and LBC, had additional cortical thinning in widespread association cortices, including bilateral TPJ
Luks et al.,	2010	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	65 mixed neurodegenerative diseases, 22 HC	MRI	-	-	-	INC attentional control	-	-	Flanker task	Voxel-based morphometry	The atrophy in the TPJ–VLPFC network was associated with slower attentional control on accurate trials
Qian et al.,	2015	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	12 aMCI, 16 AD, 15 HC	fMRI	-	-	-	-	-	-	-	Functional connectivity	Decreased functional connectivity in the orbital ventral frontal cortex and the TPJ in the AD patients, functional connectivity was preserved in aMCI compared with HC. In addition, gray matter density in the right ventral frontal cortex was correlated with functional deterioration in the rTPJ and ventral frontal cortex in AD
Rohrer et al.,	2012	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	19 AD, 23 HC	MRI	-	-	-	-	-	-	Neuropsychological testing	Cortical thickness	In patients with more severe diseases there was increased involvement of the left anterior temporal and frontal cortices and in the right hemisphere with involvement of the TPJ
Sorg et al.,	2012	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	35 AD, 23 HC for PET, 36 HC for task	PET	AIM spatial attention	-	-	-	-	-	Visual attention paradigm and partial- and whole-report paradigms	Brain metabolism	Direction and degree of spatial attention bias (towards the left or the right) was correlated with direction and degree of an interhemispheric metabolism bias in the inferior parietal lobe and the TPJ
Yamashit	2019	Neurodege	Neurodege	22 AD, 10 HC	resting-state	AIM	-	-	-	-	-	Orientation for time, a	Functional	The connectivity between the rTPJ and PCC (that is part of

a et al.,		neration	nerative disease (dementia or cognitive impairment)		fMRI	orientation for time						subscale of the MMSE	connectivity	the default mode network) was diminished in the OT-poor AD
Zamboni et al.,	2010	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	27 bvFTD, 12 aFTD, 31 CBS	MRI	-	AIM awareness and anosognosia	-	-	-	-	Anosognosia (measured as patient-caregiver discrepancy)	Voxel-based lesion-symptom mapping	The degree of anosognosia for behavioral impairment correlated with gray matter atrophy in the posterior regions of the brain, near the rTPJ
Zou et al.,	2014	Neurodegeneration	Neurodegenerative disease (dementia or cognitive impairment)	20 AD, 20 HC	MRI, magnetic resonance spectroscopy	-	-	-	-	-	-	-	Cerebral blood flow and spectrum variables	It was found a decrease in cerebral blood flow values in AD patients compared with elderly controls in the bilateral frontal regions, in the temporal lobe, in both TPJ
Hanganu et al.,	2017	Neurodegeneration	Neurodegenerative disease (Parkinson)	24 PD	MRI	-	-	-	-	-	AIM depression	Beck Depression Inventory (BDI-II) scale	Cortical thickness	Significant correlation between the worsening of depression over time and lower cortical thickness over time in the rTPJ
Pickut et al.,	2013	Neurodegeneration	Neurodegenerative disease (Parkinson)	14 PD target training, 13 PD other training	MRI	-	-	-	-	-	-	8-week mindfulness based intervention	Voxel-based morphometry	Increased gray matter density in a variety of regions including the TPJ bilaterally in the group that underwent the mindfulness training
Bosma et al.,	2018	Neurodegeneration	White matter disease	31 MS, 31 HC	Resting-state fMRI	-	-	-	-	-	-	Pain questionnaire	Static and dynamic functional connectivity	Abnormalities in the connectivity between salience network and default mode network
Carotenuto et al.,	2018	Neurodegeneration	White matter disease	36 MS	Resting-state fMRI	-	-	AIM pragmatic abilities	-	-	-	The Assessment of Pragmatic Abilities and Cognitive Substrates (APACS)	Seed-based mediated functional connectivity	Correlation between the rTPJ and deficit in pragmatic and communicative abilities
Huang et al.,	2019	Neurodegeneration	White matter disease	22 MS, 22 HC	fMRI	-	-	-	-	-	-	-	Dynamic functional connectivity	Decreased dynamic functional connectivity within both the dorsal and ventral attention networks, of which the rTPJ is part of, but increased connectivity between the two networks. Moreover, the connectivity between the rTPJ and the right ventral frontal cortex was negatively correlated with the total white matter lesion loads
Kim et al.,	2019	Neurodegeneration	White matter disease	27 MS, 26 HC	Resting-state MEG	-	-	-	-	-	AIM pain	Pain, neuropathic pain, and pain interference with activities of daily living	Resting-state regional spectral power	Increase in alpha power in the rTPJ, as part of the salience network, in particular in a subgroup of patients which experienced mixed neuropathic pain
Peng et al.,	2016	Neurodegeneration	White matter disease	23 patients with white matter hyperintense lesions, 23 HC	MRI	-	-	-	-	-	-	-	Voxel-based morphometry	Decrease and increase gray matter density in a variety of cortical regions, including the rTPJ, in which they found a general greater increase in the white matter. It was also found a negative correlation between the rTPJ and white matter lesion load hyperintensity
Beauchamp et al.,	2012	-	Other pathologies	3 epilepsy	EEG	-	-	-	-	-	-	Self-reported phoneme perception	Gamma power	Gamma oscillations in the TPJ might be a neural signature of the phosphene perception

Jiang et al.,	2018	-	Other pathologies	19 epilepsy, 21 HC	fMRI	-	-	-	-	-	-	-	Granger Causality Analysis	The influence of the default mode network was weakened in patients
Lu et al.,	2018	Post-geniculate lesions	Other pathologies	7 chronic hemianopia patients	fMRI	-	-	-	-	-	-	Contrast sensitivity	Functional connectivity	Increased activation in the rTPJ after training and a strengthening of connectivity between the anterior cingulate cortex and left insula to the rTPJ
Peterson et al.,	2018	-	Other pathologies	7 PNES	NIBS	-	-	-	-	-	-	Weekly PNES count	-	Weekly seizure rates decreased in post vs. pre-treatment. Decrease was sustained at 3-month follow-up
Zhang et al.,	2019	-	Other pathologies	29 migraine patients (26 in final analyses)	fMRI, NIBS	-	-	-	-	-	-	Disease duration, frequency of migraine attacks during the past month, score on the visual analog scale, and score on the Migraine Specific Quality-of-Life Questionnaire	Seed-to-voxel whole-brain resting-state functional connectivity	Increased functional connectivity in the rTPJ. In addition, positive association between disease duration and resting-state functional connectivity change at the rTPJ

List of acronyms (Brain areas)
lDLPFC: left dorsolateral prefrontal cortex; lPFC: lateral prefrontal cortex; lTPJ: left temporoparietal junction; PCC: posterior cingulate cortex; PFC: prefrontal cortex; rDLPFC: right dorsolateral prefrontal cortex; rTPJ: right temporoparietal junction; STG: superior temporal gyrus; VLPFC: ventrolateral prefrontal cortex.

List of acronyms (Other)
AD: Alzheimer's disease; ADCI: Alzheimer disease-related cognitive impairment; aFTD: aphasic variant frontotemporal dementia; aMCI: amnestic mild cognitive impairment; BD: patients with bilateral damage; bvFTD: behavioral variant frontotemporal dementia; CBS: corticobasal syndrome; CT: computerized tomography; DLB: dementia with Lewy bodies; EEG: electroencephalography; fMRI: functional magnetic resonance imaging; FND: functional neurological disorder; HC: healthy controls; LBCI: Lewy body disease-related cognitive impairment; LD: patients with left damage; MCI: mild cognitive impairment; MEG: magnetoencephalography; MRI: magnetic resonance imaging; MS: multiple sclerosis; NIBS: non-invasive brain stimulation; PD: Parkinson's disease; PET: positron emission tomography; PNES: Psychogenic non-epileptic seizures; proAD: prodromal Alzheimer's disease; proDLB: prodromal dementia with Lewy bodies; RD: patients with right damage; RHI: rubber hand illusion; SPECT: single photon emission computed tomography; TBI: traumatic brain injury.

Notes: For each study, the aim/s has/have been reported highlighting whether results were expected and in line with the hypotheses (“AIM” in Table) or were incidental (“INC” in Table).

4. Discussion

The current study aimed to systematically investigate the rTPJ contribution to several neurological deficits, proposing how a wide variety of deficits may be a consequence of impairment in higher-order and domain-general predictive processing, possibly but not necessarily, related to a prediction network (Siman-Tov et al., 2019). This interpretation would unify several proposals and would be in contrast to the interpretation of rTPJ as having specific roles for specific cognitive functions. Results confirmed our hypothesis that rTPJ damage or degeneration, due to direct lesion or its disconnection from other brain regions (see Table S3), has an impact on several aspects of perception, cognition, and motion. In the next paragraphs, we will argue, from a large-scale perspective, that the ubiquitous association of rTPJ with manifold functional domains is supported by widespread connections with other brain structures/networks. It will be suggested how all these domains may underlie a common domain-general prediction-related processing. Then, we will discuss the limitations of the review, as well as the implications of predictive processing on the interpretation of traditional domains of cognitive functions both for basic and clinical research.

4.1 Unveiling the ubiquity of rTPJ

Regardless of the possible bias in the literature that left some cognitive domains (e.g., language) not adequately covered, in the present review we have found a ubiquitous involvement of rTPJ across pathologies and cognitive tasks (see Table 1). In particular, it seems that damage to rTPJ can lead to impairment on almost any type of task, related to any main cognitive domain. As summarized in Table 1 and S3, findings from studies meeting the inclusion criteria showed how the damage or the disconnection of rTPJ has different functional cognitive consequences there clustered. In general, rTPJ seems crucial for attention, especially in spatial attention. The second greater cluster of deficits included studies on awareness and social cognition, considered as the inability or the difficulty of patients to maintain a proper self- or other-perception. Deficits in the representation of someone else's beliefs (i.e., ToM), difficulties to build up an appropriate body structural representation, and anosognosia are typical problems of this cluster. The systematic search shows, in addition, functional impairments involving other domains: motor deficits (including in this group also reduced goal directed behaviors for a lack of motivation, i.e., apathy) and executive functions. Finally, a marginal number of studies established that rTPJ is linked to language and memory. The sequelae of neurological disorders related to rTPJ can lead to interpreting all these deficits as the consequence of the disruption of specific domains. As such, within a modular view that assumes a one-to one correspondence between brain structures and functions, it would be reasonable to infer that rTPJ plays a key role in all these domains: attention, awareness, social cognition, executive functions, memory, language, and motor functioning. However, in the era of networks, this "local" interpretation may sound simplistic since it neglects evidence showing how these impairments can result from a variety of other etiologies (e.g., attention impairments following frontal lesions). Hence, a large-scale approach may provide a better understanding of brain functioning, also outlining a multifaceted characterization and prediction of the consequences following neurological disorders. Within a predictive framework, we argue that the ubiquitous involvement of rTPJ in cognition is supported anatomically and functionally by its connection with several brain areas, where it integrates diverse information and updates internal models and expectations regardless of the network in which it is involved. Classically, rTPJ is considered one of the core areas of the right-lateralized VAN, which controls reorienting of attention, whereas the DAN mainly controls sustained attention. Within the VAN, rTPJ processes stimuli that are unattended, but relevant to the task at hand. rTPJ is also assumed to respond to the salience of such stimuli (Corbetta et al., 2008). For example, it shows greater activation to infrequent targets in oddball tasks (Polich, 2007) and invalid rather than valid targets in a classical Posner task (Kincade et al., 2005). In a predictive view, unattended but relevant stimuli are highly salient because they violate the agent's expectations about their features (e.g., location in invalidly cued targets during the Posner cueing task), and therefore produce a greater prediction error. In line with Corbetta et al. (2008), Geng and Vossel (2013) support this view. They observed that rTPJ is active both during the oddball and the Posner cueing task. Regarding oddball tasks, TPJ is one of the most prominent neural generators of the P3, and particularly of the P3b component, which is elicited in response to deviant targets (Polich, 2007). The P3 is traditionally considered as an index of “contextual updating” reflecting the modification of the internal model of the task context based on external stimuli. Standard, frequent stimuli during oddball tasks generate strong expectations about task context, which are disconfirmed by the infrequent, deviant targets, that are highly salient and signal the need for updating the contextual expectation by integrating it with the new evidence. Geng and Vossel (2013) also interpreted the greater activation of rTPJ to the invalidly cued targets in the Posner cueing task as related to contextual updating, by a process that is analogous to that occurring for oddball tasks. These tasks activate rTPJ because this area allows the updating of the internal model of the environmental or task context based on the actual sensory inputs (Gastaldon et al., 2020). Such a change in expectation enables the initiation of actions that are appropriate to the task goal, even though during the task the stimuli have changed, or the cues are not valid. The evidence accumulated through the aforementioned studies has led some authors to hypothesize that rTPJ, besides being part of the VAN and DMN, is also part of a larger, hierarchical prediction network involved in determining the internal model of the task context (Geng and Vossel, 2013) and attention (Wilterson et al., 2021). A subsequent meta-analysis (Siman-Tov et al., 2019) supported this hypothesis, by highlighting a set of brain regions involved in both perception and action prediction, encompassing both the VAN and the DAN and including rTPJ. This reinforces the hypothesis that rTPJ is a predictive hub where different cognitive processes and information types converge and are integrated; about that, a disruption of rTPJ or its disconnection from other structures (see Table S3) - as observed in heterogeneous neurological disorders - would have consequences on several other deficits. Importantly, it seems that the role of rTPJ in predictive processing is not necessarily related to its involvement in a specific prediction network. Indeed, in such a complex and dynamic system, the exact function of rTPJ seems defined within the network which is recruited by the ongoing activity (which in experimental terms, is the ongoing task). In other words, the context drives the specific coupling between rTPJ and other regions, thus recruiting the appropriate task-specific networks. For example, if we are engaged in an attentional task (e.g., the Posner cueing task), the VAN would be recruited; if instead we are engaged in a social cognition task, the DMN would be recruited. In both cases, rTPJ would be activated always as a key hub for predictive processing, but for different purposes and together with different areas. The context- and network-dependent role of rTPJ allows for maximum flexibility of integration between different brain areas and cognitive processes. In this manner, context-dependent specific networks can use the integrative and contextual-updating function of rTPJ in different behavioral and cognitive contexts. In sum, the exact role of rTPJ is context- and network-driven, but its core function of integration and contextual updating integration does not change. We argue that the flexibility of rTPJ is the reason why we find its ubiquitous involvement across different cognitive domains: it fulfills a common function across a variety of different cognitive computations. In line with this, Fotopoulou (2014) proposed a predictive model for anosognosia for hemiplegia, reporting the existence of context-driven functional networks. This conceptualization enables us to go beyond the characterization of rTPJ as activated by peculiar, discrete, attention, or social cognition tasks contexts, and to appreciate its broader and domain-general role in comparing predictions about internal models of the environment or task to the actual external information and integrating it to the prediction themselves. This process may be impaired in neurological diseases that involve rTPJ and cause a variety of domain-general disorders. Opposite to the advocated large-scale perspective, it has been claimed that rTPJ could be subdivided into different portions with different roles, such as related to attention or ToM (Scholz et al., 2009) or to attentional selection and reorienting, a distinction supported by cytoarchitectonic differences (Gillebert et al., 2013) similarly to what has been demonstrated for the visual word form area (Weiner et al., 2017). If this very precise subdivision of rTPJ was confirmed for all the above-mentioned domains (to our knowledge, there is only limited evidence on this), one could argue that this would speak against our interpretation. Moreover, there is evidence of dissociations among the cognitive functions considered in this review, and that we traced back to damage to rTPJ or its connection to other areas (e.g., Happé et al., 1999), supporting a modular view of the mind that should be reflected in the brain architecture. We believe that both these possibilities would not alter the conclusions of the review. Considering the anatomical and functional parcellation of rTPJ, only two domains have been dissociated so far (spatial attention and ToM), while literature clearly shows the involvement of rTPJ in a plethora of cognitive and motor functions, showing the need for further studies to clarify the issue. Additionally, the proximity of two areas on two very different aspects of cognitive functions would speak in favor of the similarity of the putative role of the areas and point to a gradient-like organization of such areas (Sansom and Livesey, 2009; Vazquez-Rodríguez et al., 2019). Concerning the evidence of dissociation among the cognitive functions, although it has been largely claimed that connectionist accounts can explain dissociations even without modularity (Plaut, 1995), we would like to stress that the aim of the present review is not to challenge modularism, nor can the results from this review aim to clarify this aspect (a mixed organization of domain-specific modules and domain-general functioning would be also possible). The important point raised here is that an alternative view (based on several empirical findings) can pave new ways to design experiments both for basic research purposes and to interpret the diseases, and guide the design of new treatments, as we will suggest in the following paragraphs.

4.2 Limitations

The main limitation of the present review is that it only considers neurological populations and not healthy participants and psychiatric populations. The reason behind this choice is to be found in the aim of summarizing findings with a certain clinical relevance, to promote a predictive, more holistic view of neurological diseases. Circumscribing our work to neurological patients opens the field for future investigations on the role of rTPJ in healthy individuals. Another limit could be related to the inclusion of studies in which rTPJ was either directly lesioned or it was disconnected by other close regions (see Table S3); therefore, we cannot rule out whether the neuropsychological deficits are caused by rTPJ per se or by its disconnections with other brain areas. However, considering rTPJ as a crucial hub involved in predictive processing, damages that directly or indirectly impact rTPJ functioning are of interest to understanding the role of this area. The investigation of a single hub (like rTPJ), rather than the whole prediction network, as recently highlighted by Siman-Tov et al. (2019), can also be seen as a limitation. However, rTPJ has particular clinical relevance, as described in the Introduction. For example, it is frequently involved in right-hemisphere strokes in the territory of the middle cerebral artery and its lesioning often induces hemispatial neglect. Moreover, rTPJ is involved in cognitive functions apparently unrelated to each other (e.g., attention and ToM), suggesting the involvement of rTPJ in a plurality of networks (DAN, DMN; Corbetta et al., 2008; putative prediction network, Siman-Tov et al., 2019), which makes it susceptible to have a role in generating multiple cognitive and behavioral manifestations in neurological conditions. Finally, it is important to underline that there are two limitations related to the interpretation we made in light of a predictive processing framework. First, as already pointed out in the Introduction, there are many theoretical interpretations and perspectives on the predictive processing framework or the bayesian brain (Clark, 2013) that could lead, in the end, to different interpretations or hypotheses. Second, predictive processing implies some specific mechanisms (e.g., model updates, propagation of prediction error) and the role of rTPJ in supporting them has not been fully clarified even if it seems it could be related to the update of internal models (Geng and Vossel, 2013; Wilterson et al., 2021). Given the nature of the evidence gathered in the present review, we believe that it would be too speculative to further detail these two latter aspects. We believe the generic predictive processing framework we sketched already highlights the potential for future developments.

4.3 Implications and future directions

Considering neurological impairments engaging rTPJ as characterized by affecting domain-general predictive processing can help to shed new light on how disorders are understood. Under this perspective, the results highlighted in the current review can open new scenarios for basic research and clinical applications in neurological disorders. First, predictive processing has stimulated a paradigm shift in neuroscience that, in some cases, is leading to revising the functional organization of the brain into domains. This is affecting, for example, the distinction in those cognitive domains that have been traditionally interpreted as separate but that, under a predictive processing view, can be seen as highly connected (e.g., Doricchi et al., 2021; Fotopoulou, 2014). Nevertheless, this perspective, although not new, is still not diffuse in clinical contexts. Most clinicians involved in the assessment and rehabilitation of neurological patients still base their hypotheses and interventions on traditional domains regarding brain functioning (e.g., memory, attention, ToM), and, to our knowledge, nobody uses clinical tools based on predictive processing perspectives. Although classification in functional domains has provided a method to decompose the complexity of brain functioning in “handy” components, the firm adherence to this approach can hamper the investigation of alternative domains or modalities of functioning of the brain, with both theoretical and practical consequences. Future works can clarify whether a predictive processing framework may be advocated as a general approach that could be applied to all domains of cognition, perception, and action (i.e., domain-general) or whether there are domain-specific aspects of predictive processing. Concerning the hypothesis of domain-specific predictive mechanisms, a reinterpretation of domains related to rTPJ under the perspective of predictive processing may be plausible. For example, dedicated predictive mechanisms may be engaged in attention and be different or partially different from predictive mechanisms involved in another domain, let’s say social cognition. The domain of attention is an emblematic case of cognitive domains reinterpreted under a predictive processing framework. As mentioned, within the predictive processing framework, perception, cognition, and action occur via a continuous model updating aimed at the minimization of the prediction error. In turn, the prediction error is modulated by precision estimates that determine the weight of the prediction error, namely the brain’s degree of confidence in prediction. According to several authors, precision estimation is functionally analogous to attention (Feldman and Friston, 2010; Hohwy, 2013). Indeed, when individuals pay attention to a stimulus, an object, or a context, the reliability of the error signal is potentially higher than the opposite scenario when attention is low. Similarly, social cognition can be conceptualized as a phenomenon deriving from predictive mechanisms (see the supplementary chapter in Supplementary Materials for more details about a rethink of cognitive domains related to rTPJ under the perspective of predictive processing). However, the aim of this review is not to understand whether the brain is organized with domain-general predictive mechanisms or with domain-specific ones. Instead, regardless of the level of functional extension of these mechanisms, we want to suggest a complementary explanation for deficits related to damage or disconnection of rTPJ. Indeed, considering all of them through the lens of predictive processing may provide insights to interpret these deficits as the consequence of the disruption of predictive processing, rather than traditional modular domain-specific deficits (e.g., deficits in attention, memory, and language). We believe that bearing the more traditional hypothesis that rTPJ is involved in manifold domains raises a fundamental question: how can rTPJ contribute to all these different functions? This may be an ill-posed question because it assumes the existence of such modules. Thus, future studies might explore the general role of rTPJ as part of a prediction network supporting non-conventional domains, transcending discrete modular limits, and tending towards new ontologies (Pessoa et al., 2022; Poldrack and Yarkoni, 2016). Second, adopting a predictive processing framework can lead to very specific and clear-cut scientific questions that may further elucidate some long-lasting issues in the fields of cognitive impairments in neurological conditions. For example, it has been claimed that rTPJ may have a role in domain-general predictions, possibly as part of a prediction network (Siman-Tov et al., 2019). As Siman-Tov et al. (2019) claim, this network would be involved in higher-level predictions. But what is meant by “high-” or “low-" level of cognitive functions and ”high-” or a “low-” level of prediction? What is meant for “domain-general” (what are the domains involved)? Although labels may be sometimes fuzzy, the predictive framework can help to unveil whether traditional distinct functions have important shared aspects. For example, as recently shown by Doricchi et al. (2021), and commented by Garrido and Deouell (2021), neglect patients base their predictive behavior on statistical regularities that are related to the occurrence of sensory events on the right side of space during an auditory task; results of their study suggested how the pre-attentive reaction to the mismatch (i.e., the mismatch negativity, MMN) and the contextual updating (i.e., the P3) can be considered independent systems (Doricchi et al., 2021). The dissociation between the MMN and P3 responses may indicate that patients suffer from pre-attentive deficits, or, more specifically, low precision weights that hinder the generation of predictive processes (Garrido and Deouell, 2021; Hohwy and Seth, 2020). It would be useful to understand whether specific biomarkers could be related not just to specific domains but could be extended also to other aspects of cognition, in line with an integrative and dynamic vision of cognitive functioning (Roger and Banjac, 2022). Third, and about the First point, the interpretation of the neurological impairments in the light of a unifying approach could be a parsimonious explanation of specific deficits and could pave the way for developing innovative neuropsychological tests for assessment and new neuropsychological rehabilitation protocols (e.g., treatments). Concerning the assessment, a predictive processing view (not only related to rTPJ) supports the development of more extensive batteries to investigate clusters of deficits within and between domains (Corbetta et al., 2015), and suggests the importance to design different tasks that may be used to manipulate priors or precision and identify at what “stage” of predictive processing rTPJ plays a role. Whether the adoption of this framework has clinical relevance is an empirical issue that could be disentangled in future studies comparing the actual assessment instruments with new assessment instruments designed in light of the predictive processing perspective. Concerning the rehabilitation protocols, a predictive processing view can lead to the design of different treatments from several points of view. To give an example, the implementation of treatments in a predictive processing framework would stress the importance of improving predictive abilities in neurological patients, such as helping patients to identify and exploit probabilistic associations among sensory stimuli in the environment, or improving their ability to use contextual information to update their expectations. The idea that these deficits are not specific to individual perceptual or cognitive domains, but instead lie in mechanisms that link these domains, was previously suggested for generative models of language in schizophrenia, in which the combination of high- and low-level approaches was proposed for synergistic effects on overall cognitive and perceptual functioning (for similar considerations, see Brown and Kuperberg, 2015). Clearly, this speculative sketching of a model of predictive processing deficits in neurological disorders requires proper computational modeling and empirical testing at both neural and behavioral levels. This could have important implications for patients and for their daily functioning in a social environment, an environment that is highly complex, dynamic, and unpredictable, where an instant integration of information is required to efficiently respond to internal and external requests. In addition, even if purely speculative to date, shifting to a domaingeneral predictive processing framework may help to understand one current limitation of the traditional neuropsychological treatments: the lack of generalization of the effects or better understanding of the unexpected generalization to different domains (Anguera et al., 2013; Jacquin-Courtois et al., 2013). Assuming that it is possible to rehabilitate impaired predictive mechanisms (Brown and Kuperberg, 2015), the benefits of a treatment based on the predictive processing framework might extend over a specific (treated) domain by virtue of the fact that predictive processing is a general property of brain functioning, common across domains. As for the considerations made for assessment, it is an empirical issue that can be disentangled with experimental studies, whether this approach can lead to better and more effective rehabilitation treatments.

5. Conclusion

This review represents a systematic appraisal of studies that documented the involvement of rTPJ in neurological disorders, both in terms of its direct lesion or its disconnection from other brain regions. By overcoming the modular fragmentation of cognition following traditional domains (e.g., spatial attention, ToM, etc.), the present review sheds light on the possibility of interpreting rTPJ functioning within the overarching predicting processing framework, in particular considering rTPJ as a key hub involved in predictions for different behaviors. We suggest that the ubiquitous involvement of rTPJ in distinct domains is imputable to rTPJ being connected to several brain regions, with the functional role of integrating diverse information and updating internal models and expectations. A deeper understanding of the role of rTPJ within a prediction network paves new insights for interpreting neurological diseases and developing new treatments aimed at contrasting the impairment suffered by these patients.

References

Agosta, S., Magnago, D., Tyler, S., Grossman, E., Galante, E., Ferraro, F., Mazzini, N., Miceli, G., Battelli, L., 2017. The Pivotal Role of the Right Parietal Lobe in Temporal Attention. Journal of Cognitive Neuroscience 29, 805–815. https://doi.org/10.1162/jocn_a_01086

Aitchison, L., Lengyel, M., 2017. With or without you: predictive coding and Bayesian inference in the brain. Curr. Opin. Neurobiol., Computational Neuroscience 46, 219–227. https://doi.org/10.1016/j.conb.2017.08.010

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., Kong, E., Larraburo, Y., Rolle, C., Johnston, E., & Gazzaley, A., 2013. Video game training enhances cognitive control in older adults. *Nature*, *501*(7465), 97–101. https://doi.org/10.1038/nature12486

Arcara, G., Bambini, V., 2016. A Test for the Assessment of Pragmatic Abilities and Cognitive Substrates (APACS): Normative Data and Psychometric Properties. Frontiers in Psychology 7, 70. https://doi.org/10.3389/fpsyg.2016.00070

Baez, S., Pinasco, C., Roca, M., Ferrari, J., Couto, B., García-Cordero, I., Ibañez, A., Cruz, F., Reyes, P., Matallana, D., Manes, F., Cetcovich, M., Torralva, T., 2019. Brain structural correlates of executive and social cognition profiles in behavioral variant frontotemporal dementia and elderly bipolar disorder. Neuropsychologia, The Biological Basis of Social Cognition During Development 126, 159–169. https://doi.org/10.1016/j.neuropsychologia.2017.02.012

Bardi, L., Kanai, R., Mapelli, D., Walsh, V., 2012. TMS of the FEF Interferes with Spatial Conflict. Journal of Cognitive Neuroscience 24, 1305–1313. https://doi.org/10.1162/jocn_a_00223

Beauchamp, M.S., Sun, P., Baum, S.H., Tolias, A.S., Yoshor, D., 2012. Electrocorticography links human temporoparietal junction to visual perception. Nat Neurosci 15, 957–959. https://doi.org/10.1038/nn.3131

Blanc, F., Colloby, S.J., Philippi, N., de Pétigny, X., Jung, B., Demuynck, C., Phillipps, C., Anthony, P., Thomas, A., Bing, F., Lamy, J., Martin-Hunyadi, C., O'Brien, J.T., Cretin, B., McKeith, I., Armspach, J.-P., Taylor, J.-P., 2015. Cortical Thickness in Dementia with Lewy Bodies and Alzheimer's Disease: A Comparison of Prodromal and Dementia Stages. PLoS One 10, e0127396. https://doi.org/10.1371/journal.pone.0127396

Boccia, M., Raimo, S., Di Vita, A., Battisti, A., Matano, A., Guariglia, C., Grossi, D., Palermo, L., 2020. Topological and hodological aspects of body representation in right brain damaged patients. Neuropsychologia 148, 107637. https://doi.org/10.1016/j.neuropsychologia.2020.107637

Bosma, R.L., Kim, J.A., Cheng, J.C., Rogachov, A., Hemington, K.S., Osborne, N.R., Oh, J., Davis, K.D., 2018. Dynamic pain connectome functional connectivity and oscillations reflect multiple sclerosis pain. PAIN 159, 2267–2276. https://doi.org/10.1097/j.pain.0000000000001332

Brown, M., Kuperberg, G.R., 2015. A Hierarchical Generative Framework of Language Processing: Linking Language Perception, Interpretation, and Production Abnormalities in Schizophrenia. Front Hum Neurosci 9, 643. https://doi.org/10.3389/fnhum.2015.00643

Bzdok, D., Langner, R., Schilbach, L., Jakobs, O., Roski, C., Caspers, S., Laird, A.R., Fox, P.T., Zilles, K., Eickhoff, S.B., 2013. Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. NeuroImage 81, 381–392. https://doi.org/10.1016/j.neuroimage.2013.05.046

Campbell, B.C.V., Khatri, P., 2020. Stroke. The Lancet 396, 129–142. https://doi.org/10.1016/S0140-6736(20)31179-X

Carotenuto, A., Arcara, G., Orefice, G., Cerillo, I., Giannino, V., Rasulo, M., Iodice, R., Bambini, V., 2018. Communication in Multiple Sclerosis: Pragmatic Deficit and its Relation with Cognition and Social Cognition. Archives of Clinical Neuropsychology 33, 194–205. https://doi.org/10.1093/arclin/acx061

Carter, R.M., Huettel, S.A., 2013. A nexus model of the temporal–parietal junction. Trends in Cognitive Sciences 17, 328–336. https://doi.org/10.1016/j.tics.2013.05.007

Caspers, S., Geyer, S., Schleicher, A., Mohlberg, H., Amunts, K., Zilles, K., 2006. The human inferior parietal cortex: Cytoarchitectonic parcellation and interindividual variability. NeuroImage 33, 430–448. https://doi.org/10.1016/j.neuroimage.2006.06.054

Chechlacz, M., Rotshtein, P., Humphreys, G.W., 2014. Neuronal substrates of Corsi Block span: Lesion symptom mapping analyses in relation to attentional competition and spatial bias. Neuropsychologia 64, 240–251. https://doi.org/10.1016/j.neuropsychologia.2014.09.038

Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences 36, 181–204. https://doi.org/10.1017/S0140525X12000477

Clark, A., 2016. Surfing Uncertainty: Prediction, Action, and the Embodied Mind. Oxford University Press, New York. https://doi.org/10.1093/acprof:oso/9780190217013.001.0001

Cohen-Zimerman, S., Khilwani, H., Smith, G.N.L., Krueger, F., Gordon, B., Grafman, J., 2021. The neural basis for mental state attribution: A voxel-based lesion mapping study. Human Brain Mapping 42, 65–79. https://doi.org/10.1002/hbm.25203

Committeri, G., Piccardi, L., Galati, G., Guariglia, C., 2015. Where did you “left” Piazza del Popolo? At your “right” temporo-parietal junction. Cortex 73, 106–111. https://doi.org/10.1016/j.cortex.2015.08.009

Corbetta, M., Patel, G., Shulman, G.L., 2008. The Reorienting System of the Human Brain: From Environment to Theory of Mind. Neuron 58, 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Corbetta, M., Ramsey, L., Callejas, A., Baldassarre, A., Hacker, C.D., Siegel, J.S., Astafiev, S.V., Rengachary, J., Zinn, K., Lang, C.E., Connor, L.T., Fucetola, R., Strube, M., Carter, A.R., Shulman, G.L., 2015. Common Behavioral Clusters and Subcortical Anatomy in Stroke. Neuron 85, 927–941. https://doi.org/10.1016/j.neuron.2015.02.027

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3, 201–215. https://doi.org/10.1038/nrn755

Corbetta, M., Shulman, G.L., 2011. Spatial Neglect and Attention Networks. Annual Review of Neuroscience 34, 569–599. https://doi.org/10.1146/annurev-neuro-061010-113731

Corbetta, M., Siegel, J.S., Shulman, G.L., 2018. On the low dimensionality of behavioral deficits and alterations of brain network connectivity after focal injury. Cortex, In Memory of Professor Glyn Humphreys 107, 229–237. https://doi.org/10.1016/j.cortex.2017.12.017

De Marco, M., Ourselin, S., Venneri, A., 2019. Age and hippocampal volume predict distinct parts of default mode network activity. Sci Rep 9, 16075. https://doi.org/10.1038/s41598-019-52488-9

Decety, J., Lamm, C., 2007. The Role of the Right Temporoparietal Junction in Social Interaction: How Low-Level Computational Processes Contribute to Meta-Cognition. Neuroscientist 13, 580–593. https://doi.org/10.1177/1073858407304654

Demeurisse, G., Hublet, C., Paternot, J., Colson, C., Serniclaes, W., 1997. Pathogenesis of subcortical visuo-spatial neglect. A HMPAO SPECT study. Neuropsychologia 35, 731–735. https://doi.org/10.1016/S0028-3932(96)00126-1

Diez, I., Ortiz-Terán, L., Williams, B., Jalilianhasanpour, R., Ospina, J.P., Dickerson, B.C., Keshavan, M.S., Jr, W.C.L., Sepulcre, J., Perez, D.L., 2019. Corticolimbic fast-tracking: enhanced multimodal integration in functional neurological disorder. J Neurol Neurosurg Psychiatry 90, 929–938. https://doi.org/10.1136/jnnp-2018-319657

Doricchi, F., Pinto, M., Pellegrino, M., Marson, F., Aiello, M., Campana, S., Tomaiuolo, F., Lasaponara, S., 2021. Deficits of hierarchical predictive coding in left spatial neglect. Brain Communications 3, fcab111. https://doi.org/10.1093/braincomms/fcab111

Doya, K., Ishii, S., Pouget, A., Rao, R.P.N. (Eds.), 2006. Bayesian Brain: Probabilistic Approaches to Neural Coding, Computational Neuroscience Series. MIT Press, Cambridge, MA, USA.

Dressing, A., Martin, M., Beume, L.-A., Kuemmerer, D., Urbach, H., Kaller, C.P., Weiller, C., Rijntjes, M., 2020. The correlation between apraxia and neglect in the right hemisphere: A voxel-based lesion-symptom mapping study in 138 acute stroke patients. Cortex 132, 166–179. https://doi.org/10.1016/j.cortex.2020.07.017

Eslinger, P.J., Moore, P., Antani, S., Anderson, C., Grossman, M., 2012. Apathy in frontotemporal dementia: behavioral and neuroimaging correlates. Behav Neurol 25, 127–136. https://doi.org/10.3233/BEN-2011-0351

Euler, M.J., 2018. Intelligence and uncertainty: Implications of hierarchical predictive processing for the neuroscience of cognitive ability. Neurosci. Biobehav. Rev. 94, 93–112. https://doi.org/10.1016/j.neubiorev.2018.08.013

Feldman, H., Friston, K., 2010. Attention, Uncertainty, and Free-Energy. Frontiers in Human Neuroscience 4, 215. https://doi.org/10.3389/fnhum.2010.00215

Fodor, J.A., Fodor, P.J.A., Fodor, J., 1983. The Modularity of Mind: An Essay on Faculty Psychology. MIT Press.

Fotopoulou, A., 2014. Time to get rid of the ‘Modular’ in neuropsychology: A unified theory of anosognosia as aberrant predictive coding. Journal of Neuropsychology 8, 1–19. https://doi.org/10.1111/jnp.12010

Friston, K., 2010. The free-energy principle: a unified brain theory? Nat Rev Neurosci 11, 127–138. https://doi.org/10.1038/nrn2787

Friston, K.J., 2019. Waves of prediction. PLOS Biology 17, e3000426. https://doi.org/10.1371/journal.pbio.3000426

Frith, C. d., Wolpert, D. m., Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalizing. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 358, 459–473. https://doi.org/10.1098/rstb.2002.1218

Frucht, L., Perez, D.L., Callahan, J., MacLean, J., Song, P.C., Sharma, N., Stephen, C.D., 2021. Functional Dystonia: Differentiation From Primary Dystonia and Multidisciplinary Treatments. Frontiers in Neurology 11, 1972. https://doi.org/10.3389/fneur.2020.605262

Garrido, M.I., Deouell, L.Y., 2021. Unilateral neglect within the predictive processing framework. Brain Communications 3, fcab193. https://doi.org/10.1093/braincomms/fcab193

Gastaldon, S., Arcara, G., Navarrete, E., Peressotti, F., 2020. Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. Cortex 133, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026

Geng, J.J., Vossel, S., 2013. Re-evaluating the role of TPJ in attentional control: Contextual updating? Neuroscience & Biobehavioral Reviews 37, 2608–2620. https://doi.org/10.1016/j.neubiorev.2013.08.010

Gillebert, C. R., Mantini, D., Peeters, R., Dupont, P., & Vandenberghe, R. (2013). Cytoarchitectonic mapping of attentional selection and reorienting in parietal cortex. NeuroImage, 67, 257–272. http://dx.doi.org/10.1016/j.neuroimage.2012.11.026

Golay, L., Schnider, A., Ptak, R., 2008. Cortical and subcortical anatomy of chronic spatial neglect following vascular damage. Behavioral and Brain Functions 4, 43. https://doi.org/10.1186/1744-9081-4-43

Graziano, M.S.A., 2020. Consciousness and the attention schema: Why it has to be right. Cogn. Neuropsychol. 37, 224–233. https://doi.org/10.1080/02643294.2020.1761782

Graziano, M.S.A., Kastner, S., 2011. Human consciousness and its relationship to social neuroscience: A novel hypothesis. Cogn. Neurosci. 2, 98–113. https://doi.org/10.1080/17588928.2011.565121

Graziano, M.S.A., Webb, T.W., 2015. The attention schema theory: a mechanistic account of subjective awareness. Front. Psychol. 6.

Hanganu, A., Bruneau, M.-A., Degroot, C., Bedetti, C., Mejia-Constain, B., Lafontaine, A.-L., Chouinard, S., Monchi, O., 2017. Depressive symptoms in Parkinson's disease correlate with cortical atrophy over time. Brain and Cognition 111, 127–133. https://doi.org/10.1016/j.bandc.2016.11.001

Happé, F., Brownell, H., Winner, E., 1999. Acquired `theory of mind' impairments following stroke. Cognition 70, 211–240. https://doi.org/10.1016/S0010-0277(99)00005-0

Hattori, T., Ito, K., Nakazawa, C., Numasawa, Y., Watanabe, M., Aoki, S., Mizusawa, H., Ishiai, S., Yokota, T., 2018. Structural connectivity in spatial attention network: reconstruction from left hemispatial neglect. Brain Imaging and Behavior 12, 309–323. https://doi.org/10.1007/s11682-017-9698-7

Heeger, D.J., 2017. Theory of cortical function. Proc. Natl. Acad. Sci. 114, 1773–1782. https://doi.org/10.1073/pnas.1619788114

Hohwy, J., 2013. The Predictive Mind. Oxford University Press.

Hohwy, J., 2017. Priors in perception: Top-down modulation, Bayesian perceptual learning rate, and prediction error minimization. Conscious. Cogn. 47, 75–85. https://doi.org/10.1016/j.concog.2016.09.004

Hohwy, J., Seth, A., 2020. Predictive processing as a systematic basis for identifying the neural correlates of consciousness. Philosophy and the Mind Sciences 1. https://doi.org/10.33735/phimisci.2020.II.64

Howes, O.D., Hird, E.J., Adams, R.A., Corlett, P.R., McGuire, P., 2020. Aberrant Salience, Information Processing, and Dopaminergic Signaling in People at Clinical High Risk for Psychosis. Biol. Psychiatry, Psychosis Risk Syndrome 88, 304–314. https://doi.org/10.1016/j.biopsych.2020.03.012

Huang, M., Zhou, F., Wu, L., Wang, B., Guo, L., Zhao, Y., Wan, H., Li, F., Zeng, X., Gong, H., 2019. White matter lesion loads associated with dynamic functional connectivity within attention network in patients with relapsing-remitting multiple sclerosis. Journal of Clinical Neuroscience 65, 59–65. https://doi.org/10.1016/j.jocn.2019.03.034

Jacquin-Courtois, S., O'Shea, J., Luauté, J., Pisella, L., Revol, P., Mizuno, K., Rode, G., & Rossetti, Y., 2013. Rehabilitation of spatial neglect by prism adaptation. A peculiar expansion of sensorimotor after-effects to spatial cognition. *Neuroscience and Biobehavioral Reviews*, *37*(4), 594–609. https://doi.org/10.1016/j.neubiorev.2013.02.007

Jiang, L.-W., Qian, R.-B., Fu, X.-M., Zhang, D., Peng, N., Niu, C.-S., Wang, Y.-H., 2018. Altered attention networks and DMN in refractory epilepsy: A resting-state functional and causal connectivity study. Epilepsy & Behavior 88, 81–86. https://doi.org/10.1016/j.yebeh.2018.06.045

Kang, S.W., Jeon, S., Yoo, H.S., Chung, S.J., Lee, P.H., Sohn, Y.H., Yun, M., Evans, A.C., Ye, B.S., 2019. Effects of Lewy body disease and Alzheimer disease on brain atrophy and cognitive dysfunction. Neurology 92, e2015–e2026. https://doi.org/10.1212/WNL.0000000000007373

Karnath, H.-O., Himmelbach, M., Küker, W., 2003. The cortical substrate of visual extinction. NeuroReport 14, 437–442.

Kaski, D., Quadir, S., Nigmatullina, Y., Malhotra, P.A., Bronstein, A.M., Seemungal, B.M., 2016. Temporoparietal encoding of space and time during vestibular-guided orientation. Brain 139, 392–403. https://doi.org/10.1093/brain/awv370

Kaufman, A., Serfaty, C., Deouell, L.Y., Ruppin, E., Soroker, N., 2009. Multiperturbation analysis of distributed neural networks: the case of spatial neglect. Hum Brain Mapp 30, 3687–3695. https://doi.org/10.1002/hbm.20797

Keller, G.B., Mrsic-Flogel, T.D., 2018. Predictive Processing: A Canonical Cortical Computation. Neuron 100, 424–435. https://doi.org/10.1016/j.neuron.2018.10.003

Kim, J.A., Bosma, R.L., Hemington, K.S., Rogachov, A., Osborne, N.R., Cheng, J.C., Oh, J., Crawley, A.P., Dunkley, B.T., Davis, K.D., 2019. Neuropathic pain and pain interference are linked to alpha-band slowing and reduced beta-band magnetoencephalography activity within the dynamic pain connectome in patients with multiple sclerosis. PAIN 160, 187–197. https://doi.org/10.1097/j.pain.0000000000001391

Kincade, J.M., Abrams, R.A., Astafiev, S.V., Shulman, G.L., Corbetta, M., 2005. An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. J Neurosci 25, 4593–4604. https://doi.org/10.1523/JNEUROSCI.0236-05.2005

Kocagoncu, E., Klimovich-Gray, A., Hughes, L.E., Rowe, J.B., 2021. Evidence and implications of abnormal predictive coding in dementia. Brain J. Neurol. 144, 3311–3321. https://doi.org/10.1093/brain/awab254

Koster-Hale, J., Saxe, R., 2013. Theory of Mind: A Neural Prediction Problem. Neuron 79, 836–848. https://doi.org/10.1016/j.neuron.2013.08.020

Krall, S.C., Rottschy, C., Oberwelland, E., Bzdok, D., Fox, P.T., Eickhoff, S.B., Fink, G.R., Konrad, K., 2015. The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. Brain Struct Funct 220, 587–604. https://doi.org/10.1007/s00429-014-0803-z

Kube, T., Schwarting, R., Rozenkrantz, L., Glombiewski, J.A., Rief, W., 2020. Distorted Cognitive Processes in Major Depression: A Predictive Processing Perspective. Biol. Psychiatry, Mechanisms of Major Depression 87, 388–398. https://doi.org/10.1016/j.biopsych.2019.07.017

Kubit, B., Jack, A., 2013. Rethinking the role of the rTPJ in attention and social cognition in light of the opposing domains hypothesis: findings from an ALE-based meta-analysis and resting-state functional connectivity. Frontiers in Human Neuroscience 7, 323. https://doi.org/10.3389/fnhum.2013.00323

Lee, B.H., Suh, M.K., Kim, E.-J., Seo, S.W., Choi, K.M., Kim, G.-M., Chung, C.-S., Heilman, K.M., Na, D.L., 2009. Neglect dyslexia: Frequency, association with other hemispatial neglects, and lesion localization. Neuropsychologia 47, 704–710. https://doi.org/10.1016/j.neuropsychologia.2008.11.027

Leigh, R., Oishi, K., Hsu, J., Lindquist, M., Gottesman, R.F., Jarso, S., Crainiceanu, C., Mori, S., Hillis, A.E., 2013. Acute lesions that impair affective empathy. Brain 136, 2539–2549. https://doi.org/10.1093/brain/awt177

Lu, Q., Wang, X., Li, L., Qiu, B., Wei, S., Sabel, B.A., Zhou, Y., 2018. Visual rehabilitation training alters attentional networks in hemianopia: An fMRI study. Clinical Neurophysiology 129, 1832–1841. https://doi.org/10.1016/j.clinph.2018.05.027

Luks, T.L., Oliveira, M., Possin, K.L., Bird, A., Miller, B.L., Weiner, M.W., Kramer, J.H., 2010. Atrophy in two attention networks is associated with performance on a Flanker task in neurodegenerative disease. Neuropsychologia 48, 165–170. https://doi.org/10.1016/j.neuropsychologia.2009.09.001

Mandonnet, E., Cerliani, L., Siuda-Krzywicka, K., Poisson, I., Zhi, N., Volle, E., de Schotten, M.T., 2017. A network-level approach of cognitive flexibility impairment after surgery of a right temporo-parietal glioma. Neurochirurgie 63, 308–313. https://doi.org/10.1016/j.neuchi.2017.03.003

Martinaud, O., Besharati, S., Jenkinson, P.M., Fotopoulou, A., 2017. Ownership illusions in patients with body delusions: Different neural profiles of visual capture and disownership. Cortex, Confabulation and related disorders 87, 174–185. https://doi.org/10.1016/j.cortex.2016.09.025

Mengotti, P., Dombert, P.L., Fink, G.R., Vossel, S., 2017. Disruption of the Right Temporoparietal Junction Impairs Probabilistic Belief Updating. J Neurosci 37, 5419–5428. https://doi.org/10.1523/JNEUROSCI.3683-16.2017

Monai, E., Bernocchi, F., Bisio, M., Bisogno, A.L., Salvalaggio, A., Corbetta, M., 2020. Multiple Network Disconnection in Anosognosia for Hemiplegia. Frontiers in Systems Neuroscience 14, 21. https://doi.org/10.3389/fnsys.2020.00021

Nave, K., Deane, G., Miller, M., Clark, A., 2020. Wilding the predictive brain. WIREs Cogn. Sci. 11, e1542. https://doi.org/10.1002/wcs.1542

Ng, Y.S., Stein, J., Ning, M., Black-Schaffer, R.M., 2007. Comparison of Clinical Characteristics and Functional Outcomes of Ischemic Stroke in Different Vascular Territories. Stroke 38, 2309–2314. https://doi.org/10.1161/STROKEAHA.106.475483

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372, n71. https://doi.org/10.1136/bmj.n71

Palop, J.J., Chin, J., Mucke, L., 2006. A network dysfunction perspective on neurodegenerative diseases. Nature 443, 768–773. https://doi.org/10.1038/nature05289

Park, B., Fareri, D., Delgado, M., Young, L., 2021. The role of right temporoparietal junction in processing social prediction error across relationship contexts. Social Cognitive and Affective Neuroscience 16, 772–781. https://doi.org/10.1093/scan/nsaa072

Patel, G.H., Sestieri, C., Corbetta, M., 2019. The evolution of the temporoparietal junction and posterior superior temporal sulcus. Cortex 118, 38–50. https://doi.org/10.1016/j.cortex.2019.01.026

Patra, J., Bhatia, M., Suraweera, W., Morris, S.K., Patra, C., Gupta, P.C., Jha, P., 2015. Exposure to Second-Hand Smoke and the Risk of Tuberculosis in Children and Adults: A Systematic Review and Meta-Analysis of 18 Observational Studies. PLOS Medicine 12, e1001835. https://doi.org/10.1371/journal.pmed.1001835

Pedrazzini, E., Ptak, R., 2019. Damage to the right temporoparietal junction, but not lateral prefrontal or insular cortex, amplifies the role of goal-directed attention. Sci Rep 9, 306. https://doi.org/10.1038/s41598-018-36537-3

Pedrazzini, E., Ptak, R., 2020. The neuroanatomy of spatial awareness: a large-scale region-of-interest and voxel-based anatomical study. Brain Imaging Behav 14, 615–626. https://doi.org/10.1007/s11682-019-00213-5

Pedrazzini, E., Schnider, A., Ptak, R., 2017. A neuroanatomical model of space-based and object-centered processing in spatial neglect. Brain Struct Funct 222, 3605–3613. https://doi.org/10.1007/s00429-017-1420-4

Peng, Y., Li, S., Zhuang, Y., Liu, X., Wu, L., Gong, H., Liu, D., Zhou, F., 2016. Density abnormalities in normal-appearing gray matter in the middle-aged brain with white matter hyperintense lesions: a DARTEL-enhanced voxel-based morphometry study. Clin Interv Aging 11, 615–622. https://doi.org/10.2147/CIA.S98409

Pessoa, L., Medina, L., Desfilis, E., 2022. Refocusing neuroscience: moving away from mental categories and towards complex behaviours. Philosophical Transactions of the Royal Society B: Biological Sciences 377, 20200534. https://doi.org/10.1098/rstb.2020.0534

Peterson, K.T., Kosior, R., Meek, B.P., Ng, M., Perez, D.L., Modirrousta, M., 2018. Right Temporoparietal Junction Transcranial Magnetic Stimulation in the Treatment of Psychogenic Non-epileptic Seizures: A Case Series. Psychosomatics 59, 601–606. https://doi.org/10.1016/j.psym.2018.03.001

Pezzulo, G., Zorzi, M., Corbetta, M., 2021. The secret life of predictive brains: what's spontaneous activity for? Trends in Cognitive Sciences 25, 730–743. https://doi.org/10.1016/j.tics.2021.05.007

Pickut, B.A., Van Hecke, W., Kerckhofs, E., Mariën, P., Vanneste, S., Cras, P., Parizel, P.M., 2013. Mindfulness based intervention in Parkinson's disease leads to structural brain changes on MRI: A randomized controlled longitudinal trial. Clinical Neurology and Neurosurgery 115, 2419–2425. https://doi.org/10.1016/j.clineuro.2013.10.002

Pisella, L., Alahyane, N., Blangero, A., Thery, F., Blanc, S., Pelisson, D., 2011. Right-hemispheric dominance for visual remapping in humans. Philosophical Transactions of the Royal Society B: Biological Sciences 366, 572–585. https://doi.org/10.1098/rstb.2010.0258

Plaut, D.C., 1995. Double dissociation without modularity: Evidence from connectionist neuropsychology. J. Clin. Exp. Neuropsychol. 17, 291–321. https://doi.org/10.1080/01688639508405124

Poldrack, R.A., Yarkoni, T., 2016. From brain maps to cognitive ontologies: informatics and the search for mental structure. Annu Rev Psychol 67, 587–612. https://doi.org/10.1146/annurev-psych-122414-033729

Polich, J., 2007. Updating P300: An integrative theory of P3a and P3b. Clinical Neurophysiology 118, 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Ptak, R., Schnider, A., 2011. The attention network of the human brain: Relating structural damage associated with spatial neglect to functional imaging correlates of spatial attention. Neuropsychologia 49, 3063–3070. https://doi.org/10.1016/j.neuropsychologia.2011.07.008

Qian, S., Zhang, Z., Li, B., Sun, G., 2015. Functional-structural degeneration in dorsal and ventral attention systems for Alzheimer's disease, amnestic mild cognitive impairment. Brain Imaging Behav 9, 790–800. https://doi.org/10.1007/s11682-014-9336-6

Rahnev, D., 2019. The Bayesian brain: What is it and do humans have it? Behav. Brain Sci. 42. https://doi.org/10.1017/S0140525X19001377

Roger, E., Banjac, S., Thiebaut de Schotten, M., Baciu, M., 2022. Missing links: The functional unification of language and memory (LUM). Neuroscience & Biobehavioral Reviews 133, 104489. https://doi.org/10.1016/j.neubiorev.2021.12.012

Rohrer, J.D., Rossor, M.N., Warren, J.D., 2012. Alzheimer's pathology in primary progressive aphasia. Neurobiol Aging 33, 744–752. https://doi.org/10.1016/j.neurobiolaging.2010.05.020

Rousseaux, M., Allart, E., Bernati, T., Saj, A., 2015. Anatomical and psychometric relationships of behavioral neglect in daily living. Neuropsychologia 70, 64–70. https://doi.org/10.1016/j.neuropsychologia.2015.02.011

Sanborn, A.N., Chater, N., 2016. Bayesian Brains without Probabilities. Trends Cogn. Sci. 20, 883–893. https://doi.org/10.1016/j.tics.2016.10.003

Sansom, S.N., Livesey, F.J., 2009. Gradients in the brain: the control of the development of form and function in the cerebral cortex. Cold Spring Harb. Perspect. Biol. 1, a002519. https://doi.org/10.1101/cshperspect.a002519

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E.N., Saxe, R., 2009. Distinct Regions of Right Temporo-Parietal Junction Are Selective for Theory of Mind and Exogenous Attention. PLOS ONE 4, e4869. https://doi.org/10.1371/journal.pone.0004869

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J., 2014. Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, *42*, 9–34. https://doi.org/10.1016/j.neubiorev.2014.01.009

Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P., 2021. Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, *147*(3), 293–327. https://doi.org/10.1037/bul0000303

Schurz, M., Tholen, M. G., Perner, J., Mars, R. B., & Sallet, J., 2017. Specifying the brain anatomy underlying temporo-parietal junction activations for theory of mind: A review using probabilistic atlases from different imaging modalities. *Human Brain Mapping*, *38*(9), 4788–4805. https://doi.org/10.1002/hbm.23675

Shallice, T., 1988. From Neuropsychology to Mental Structure. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511526817

Shomstein, S., Lee, J., Behrmann, M., 2010. Top-down and bottom-up attentional guidance: investigating the role of the dorsal and ventral parietal cortices. Exp Brain Res 206, 197–208. https://doi.org/10.1007/s00221-010-2326-z

Siman-Tov, T., Granot, R.Y., Shany, O., Singer, N., Hendler, T., Gordon, C.R., 2019. Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction. Neurosci Biobehav Rev 105, 262–275. https://doi.org/10.1016/j.neubiorev.2019.08.012

Singh, J., Knight, R.T., 1993. Effects of posterior association cortex lesions on brain potentials preceding self-initiated movements. J. Neurosci. 13, 1820–1829. https://doi.org/10.1523/JNEUROSCI.13-05-01820.1993

Smith, R., Badcock, P., Friston, K.J., 2021. Recent advances in the application of predictive coding and active inference models within clinical neuroscience. Psychiatry and Clinical Neurosciences 75, 3–13. https://doi.org/10.1111/pcn.13138

Sorg, C., Myers, N., Redel, P., Bublak, P., Riedl, V., Manoliu, A., Perneczky, R., Grimmer, T., Kurz, A., Förstl, H., Drzezga, A., Müller, H.J., Wohlschläger, A.M., Finke, K., 2012. Asymmetric loss of parietal activity causes spatial bias in prodromal and mild Alzheimer’s disease. Biol Psychiatry 71, 798–804. https://doi.org/10.1016/j.biopsych.2011.09.027

Starkstein, S.E., Fedoroff, J.P., Price, T.R., Leiguarda, R., Robinson, R.G., 1992. Anosognosia in patients with cerebrovascular lesions. A study of causative factors. Stroke 23, 1446–1453. https://doi.org/10.1161/01.STR.23.10.1446

Swanson, L.R., 2016. The Predictive Processing Paradigm Has Roots in Kant. Frontiers in Systems Neuroscience 10, 79. https://doi.org/10.3389/fnsys.2016.00079

Thiebaut de Schotten, M., Tomaiuolo, F., Aiello, M., Merola, S., Silvetti, M., Lecce, F., Bartolomeo, P., Doricchi, F., 2014. Damage to White Matter Pathways in Subacute and Chronic Spatial Neglect: A Group Study and 2 Single-Case Studies with Complete Virtual "In Vivo" Tractography Dissection. Cerebral Cortex 24, 691–706. https://doi.org/10.1093/cercor/bhs351

Thompson, K.G., Hanes, D.P., Bichot, N.P., Schall, J.D., 1996. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. Journal of Neurophysiology 76, 4040–4055. https://doi.org/10.1152/jn.1996.76.6.4040

Ticini, L.F., de Haan, B., Klose, U., Nägele, T., Karnath, H.-O., 2010. The Role of Temporo-parietal Cortex in Subcortical Visual Extinction. Journal of Cognitive Neuroscience 22, 2141–2150. https://doi.org/10.1162/jocn.2009.21315

Toba, M.N., Zavaglia, M., Malherbe, C., Moreau, T., Rastelli, F., Kaglik, A., Valabrègue, R., Pradat-Diehl, P., Hilgetag, C.C., Valero-Cabré, A., 2020. Game theoretical mapping of white matter contributions to visuospatial attention in stroke patients with hemineglect. Hum Brain Mapp 41, 2926–2950. https://doi.org/10.1002/hbm.24987

Vallar, G., 2007. Spatial Neglect, Balint-Homes' and Gerstmann's Syndrome, and Other Spatial Disorders. CNS Spectrums 12, 527–536. https://doi.org/10.1017/S1092852900021271

Vallar, G., Calzolari, E., 2018. Chapter 14 - Unilateral spatial neglect after posterior parietal damage, in: Vallar, G., Coslett, H.B. (Eds.), Handbook of Clinical Neurology, The Parietal Lobe. Elsevier, pp. 287–312. https://doi.org/10.1016/B978-0-444-63622-5.00014-0

van Elk, M., 2021. A predictive processing framework of tool use. Cortex 139, 211–221. https://doi.org/10.1016/j.cortex.2021.03.014

Vázquez-Rodríguez, B., Suárez, L.E., Markello, R.D., Shafiei, G., Paquola, C., Hagmann, P., Heuvel, M.P. van den, Bernhardt, B.C., Spreng, R.N., Misic, B., 2019. Gradients of structure–function tethering across neocortex. Proc. Natl. Acad. Sci. 116, 21219–21227. https://doi.org/10.1073/pnas.1903403116

Verleger, R., Heide, W., Butt, C., Kömpf, D., 1994. Reduction of P3b in patients with temporo-parietal lesions. Cognitive Brain Research 2, 103–116. https://doi.org/10.1016/0926-6410(94)90007-8

Vossel, S., Geng, J.J., Fink, G.R., 2014. Dorsal and Ventral Attention Systems: Distinct Neural Circuits but Collaborative Roles. Neuroscientist 20, 150–159. https://doi.org/10.1177/1073858413494269

Vossel, S., Thiel, C.M., Fink, G.R., 2006. Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. NeuroImage 32, 1257–1264. https://doi.org/10.1016/j.neuroimage.2006.05.019

Wawrzyniak, M., Klingbeil, J., Zeller, D., Saur, D., Classen, J., 2018. The neuronal network involved in self-attribution of an artificial hand: A lesion network-symptom-mapping study. NeuroImage 166, 317–324. https://doi.org/10.1016/j.neuroimage.2017.11.011

Weiner, K. S., Barnett, M. A., Lorenz, S., Caspers, J., Stigliani, A., Amunts, K., Zilles, K., Fischl, B., & Grill-Spector, K. (2017). The Cytoarchitecture of Domain-specific Regions in Human High-level Visual Cortex. Cerebral Cortex, 27, 146–161. https://doi.org/10.1093/cercor/bhw361

Wells, G., O'Connell, D., Peterson, J., Welch, W., Losos, M., Tugwell, P., 2011. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses.

Wilson, B., Winegardner, J., Heugten, C. van, Ownsworth, T., 2017. Neuropsychological Rehabilitation: The International Handbook, 1° edizione. ed. Routledge, London ; New York.

Wilterson, A.I., Nastase, S.A., Bio, B.J., Guterstam, A., Graziano, M.S.A., 2021. Attention, awareness, and the right temporoparietal junction. Proc. Natl. Acad. Sci. 118, e2026099118. https://doi.org/10.1073/pnas.2026099118

Yamashita, K.-I., Uehara, T., Prawiroharjo, P., Yamashita, K., Togao, O., Hiwatashi, A., Taniwaki, Y., Utsunomiya, H., Matsushita, T., Yamasaki, R., Kira, J.-I., 2019. Functional connectivity change between posterior cingulate cortex and ventral attention network relates to the impairment of orientation for time in Alzheimer’s disease patients. Brain Imaging Behav 13, 154–161. https://doi.org/10.1007/s11682-018-9860-x

Yon, D., Frith, C.D., 2021. Precision and the Bayesian brain. Curr. Biol. 31, R1026–R1032. https://doi.org/10.1016/j.cub.2021.07.044

Zamboni, G., Grafman, J., Krueger, F., Knutson, K.M., Huey, E.D., 2010. Anosognosia for behavioral disturbances in frontotemporal dementia and corticobasal syndrome: A voxel-based morphometry study. Dement Geriatr Cogn Disord 29, 88–96. https://doi.org/10.1159/000255141

Zhang, Y., Liu, J., Li, H., Yan, Z., Liu, X., Cao, J., Park, J., Wilson, G., Liu, B., Kong, J., 2019. Transcutaneous auricular vagus nerve stimulation at 1 Hz modulates locus coeruleus activity and resting state functional connectivity in patients with migraine: An fMRI study. NeuroImage: Clinical 24, 101971. https://doi.org/10.1016/j.nicl.2019.101971

Zou, J.-X., Wang, M.-J., Lei, X.-J., Chen, X.-G., 2014. 3.0T MRI arterial spin labeling and magnetic resonance spectroscopy technology in the application of Alzheimer’s disease. Experimental Gerontology 60, 31–36. https://doi.org/10.1016/j.exger.2014.09.009

Interim summary 1

The first study focused on rTPJ (as it represent most widely investigated area among TPJs) to show its potentially domain-general role. Study 2 will momentarily shift the attention away from the TPJs, as we will delve more in-depth into linguistic predictive dynamics, by investigating the unfolding of the neurophysiological correlates of linguistic and semantic prediction generation and violation.

Prediction is known to take place during language (e.g., Gastaldon et al., 2020; Nicenboim et al., 2020; Nieuwland et al., 2020), but it has been mostly investigated with violation paradigms, which only consider what happens after predictions have been (dis)confirmed (similarly to those studies on the predictive function of the TPJs, that detect activations in conditions of prediction violation). Linguistic violation paradigms often elicit N400 and late posterior positivity/P600 ERPs, which have been interestingly linked to prediction error and subsequent repair or reanalysis processes, respectively (Kutas & Hillyard, 1984; Van De Meerendonk et al., 2010). In addition, it has been recently hypothesized that the N400 and the late posterior positivity/P600 could share some similarities with ERPs that are classically found in perceptual tasks, i.e., the mismatch negativity (MMN) and P300 respectively, and that their increased latencies and different topographies reflect the increased complexity of linguistic stimuli when compared to perceptual ones (Bornkessel-Schlesewski and Schlesewski, 2019; Leckey & Federmeier, 2020). Interestingly, one of the neural generators of the P300 is the TPJ (Polich, 2007); therefore, if the late posterior positivity/P600 really shares some similarities with the P300, it would be worth to explore the possibility that neurophysiological correlates of linguistic prediction can be further linked to the TPJs. We will attempt this in Study 3, and from this moment on, this work will mainly focus on prediction in the language domain and its relation with TPJs activity.

Study 2: Trial-by-trial fluctuations of pre-stimulus alpha power predict language ERPs

This study was published in Psychophysiology: Lago, S., Pezzetta, R., Gastaldon, S., Peressotti, F., & Arcara, G. (2023). Trial-by-trial fluctuations of pre-stimulus alpha power predict language ERPs. *Psychophysiology*, 00:e14388. https://doi.org/10.1111/psyp.14388

1. Introduction

1.1. Anticipation in language: ERPs and pre-stimulus alpha

Electrophysiology (EEG), and especially event-related potentials (ERPs), have long been used to study neurocognitive correlates of language processing (Bornkessel-Schlesewsky & Schlesewsky, 2019). Some of the most studied ERPs include the N400 but also post-N400 positivities such as the late posterior positivity/P600, which are elicited in linguistic violation paradigms where the final target word of a sentence is incongruent with the preceding context. Traditionally, the N400 and the late posterior positivity/P600 have been associated with incongruent sentence endings and reanalysis of stimuli perceived as anomalous, respectively (Kutas & Hillyard, 1984; Van De Meerendonk et al., 2010); recently, however, these modulations have been interpreted according to predictive processing theories (Kuperberg et al., 2020; Mantegna et al., 2019) that propose unifying frameworks for understanding neural computations underlying perception, cognition, and action (Clark, 2013; Friston, 2012). In predictive processing frameworks, to deal with environmental changes, the brain predicts upcoming events through the implementation of internal models that spontaneously infer causal structures in the world (Clark, 2013). An internal model identifies causal regularities from a complex variety of sensory signals, extracting what is relevant to predict future stimuli, and detects mismatches between predictions (priors) and actual sensory inputs. The mismatch between a prediction and the actual sensory input generates a prediction error, which is used to update expectations and provide better predictions that are, in turn, tested against the actual input (Friston, 2019). According to some predictive processing perspectives, the N400 could reflect prediction errors, i.e., the degree to which the linguistic stimulus deviates from predictions generated from the context (e.g., Hodapp & Rabovski, 2021), even though some studies have argued that such an effect can be explained without considering top-down prediction[2] (Huettig & Mani, 2016). On the other hand, the late posterior positivity/P600 is associated with a variety of tasks where linguistic stimuli are perceived as syntactically or semantically anomalous (Quante et al., 2018; Van Petten & Luka, 2012) and would reflect the attempt at resolving prediction error through reanalysis of the stimulus and the preceding context (Kuperberg et al., 2020; Wang et al., 2021). Albeit there exists compelling evidence from a wide range of studies that anticipation indeed occurs during language comprehension and production (e.g., Gastaldon et al., 2020; Nicenboim et al., 2020; Nieuwland et al., 2020), linguistic violation paradigms have the main limitation of investigating only what happens after predictions have been disconfirmed (or confirmed), thereby assuming that anticipation is taking place before the target, but not directly addressing it (León-Cabrera et al., 2017). To fill this gap, some recent studies have investigated how expectations modulate neurophysiological dynamics in the pre-stimulus interval, i.e., before the target word. In these studies, participants read sentences that led them to make strong vs weak predictions about the upcoming target word (the sentence contexts differed in their level of constraint). As compared to the post-stimulus activity (typically focused on ERP components such as N400 or P600), investigations on pre-stimulus activity are typically focused on frequency-specific signals in the brain, such as modulations of alpha (8-13 Hz) and beta frequencies (14-30 Hz). In particular, it is well-known that decreases in alpha power are associated with increased attention and stimulus processing, but also with expectations about the upcoming stimuli (see Section 1.2.).

---

[2] Note that predictive processing has been also associated with other neurophysiological correlates (i.e., oscillatory activity in specific frequency bands, Lewis & Bastiaansen, 2015). The present article focuses on the traditional language-related ERP components.

Contrasts between different levels of contextual constraint indicated that expectations modulate alpha power in the pre-stimulus interval: stronger predictions are associated with lower levels of pre-stimulus alpha power than sentences leading to weaker predictions (Gastaldon et al., 2020; León-Cabrera et al., 2022; Rommers et al., 2017; Terporten et al., 2019; Wang et al., 2018). Some of these studies also investigated N400 amplitude in relation to the preceding context and found that contexts eliciting lower pre-stimulus alpha power are associated with more negative post-stimulus N400s (Gastaldon et al., 2023; Rommers et al., 2017; Terporten et al., 2019; Wang et al., 2018). Even though the current evidence points to a link between pre-stimulus alpha and N400 amplitude, to our knowledge post-N400 positivities, such as the late posterior positivity/P600, have not been investigated yet in relation to pre-stimulus dynamics. Also, albeit a reduction in pre-stimulus alpha power has been linked to linguistic expectation, it should be kept in mind that the studies reviewed above have contrasted alpha activity across conditions; therefore, these results are relative and are not informative on how pre-stimulus alpha *per se* and regardless of experimental manipulation may modulate post-stimulus responses (Van Diepen et al., 2019).

1.2. Spontaneous fluctuations in alpha power: attention and precision

Fluctuations in pre-stimulus alpha power can occur spontaneously, i.e., in absence of experimental manipulations. Spontaneous, trial-by-trial fluctuations in pre-stimulus alpha power have been especially studied in the attention and perception domains, revealing that in visual tasks, lower pre-stimulus alpha power is linked to conscious perception of brief and unexpected visual stimuli (Ergenoglu et al., 2004; Hanslmayr et al., 2007; Hutchinson et al., 2021; Limbach & Corballis, 2016) or to shorter reaction times (Min & Herrmann, 2007; Samaha et al., 2018; Van Den Berg et al., 2016). Besides behavioral performance, variations in pre-stimulus alpha power seem to be related also to changes in latency and amplitude of post-stimulus ERPs: lower levels of pre-stimulus alpha power have been associated with larger P1, N1, P3, and P3b amplitudes (Cao et al., 2017, Ergenoglu et al., 2004; Grent-'T-Jong et al., 2011; Samaha et al., 2018; Van Den Berg et al., 2016), while higher pre-stimulus alpha power was associated with the suppression of early ERP components (C1 and N150), longer P3 link latencies and enhancement of late ERPs (after 400 msec; Iemi et al., 2019; Min & Herrmann, 2007).

Findings from these studies seem to support theories that conceptualize alpha oscillations as an anticipatory attentional suppression mechanism (Foxe & Snyder, 2011) or as a gate for information to be redirected to task-relevant brain regions, while task-irrelevant ones are inhibited (Jensen & Mazaheri, 2010). According to these views, anticipatory increases in alpha power reflect a state of cortical inhibition, while anticipatory alpha power reductions reflect cortical activation and facilitate the subsequent detection and processing of task-relevant information (Klimesch et al., 2007). Such a line of reasoning could also explain the above-mentioned association between low pre-stimulus alpha power and larger N400s in linguistic paradigms: in a state of anticipatory attention or cortical activation, the processing of linguistic stimuli is facilitated, and so will be the processing of incongruent targets that elicit N400s.

In the context of predictive processing, alpha oscillations have been interpreted in different ways (Alamia & VanRullen, 2019; Friston, 2019). Some studies link specific properties of pre-stimulus alpha to top-down predictions (Alamia & Van Rullen, 2019; Cao et al., 2017), and especially to the concept of precision (Bauer et al., 2014; Sedley et al., 2016; Sherman et al., 2016). Precision is defined mathematically as the inverse variance of a prediction and represents its reliability (Friston, 2018), but it can also be associated with perceived stimuli or prediction errors (Walsh et al., 2020). In sum, the more precise a stimulus or a prediction, the more the individual will rely on it.

Interestingly, attention has been considered as an emergent property of the precision optimization mechanism that takes place during prediction, where attending to a stimulus means representing and increasing the precision of sensory information (and prediction error) during the inferential process (Feldman & Friston, 2010). This is consistent with the results of studies reporting larger neural responses to attended vs. unattended stimuli (e.g., Jiang et al., 2013), but although this finding is also in line with the traditional accounts of attention, predictive processing is different from these accounts because it proposes a constant interaction between attention and expectation (Walsh et al., 2020).

To illustrate attentional dynamics in light of predictive processing, Feldman and Friston (2010) simulated the evoked responses to a classical Posner paradigm, where the cue preceding the target could be valid (correctly indicating the position where the target will appear) or invalid (incorrectly indicating the target position). The cue, either valid or invalid, modulates the precision of the predictions about the target before it appears because it induces a highly precise expectation, thus reducing the uncertainty about the upcoming stimulus. When the stimulus appears, the resulting prediction error precision is increased, enhancing the stimulus-evoked responses. In this way, attentional focus results from the predictability of the stimulus (attentional resources will be biased towards more precise predictions; Hohwy, 2012), and the magnitude of expectations' effects on neural activity depends on the precision-weighting applied to particular error signals (Walsh et al., 2020). However, the cues' characteristics do not map linearly with their internal representations' precision, i.e., with the predictions' precision about the target: generative models include random variations of precision (state-dependent error variance) and assume that it is not constant for any level of hierarchical inference (Feldman & Friston, 2010). This means that precision might fluctuate spontaneously during the task, and points to the importance of studying such spontaneous fluctuations of precision during predictive processing, which is precisely the scope of this paper.

To sum up, attention depends in part on the precision (i.e., reliability) of the prior (Hohwy, 2012), and highly precise (reliable) predictions seem to be associated with lower pre-stimulus alpha power in perceptual paradigms (Bauer et al., 2014; Cao et al., 2017), while higher levels of pre-stimulus alpha power have been associated with attenuated stimulus-evoked responses (Cao et al., 2017).

This view is in line with the above-mentioned findings on the role of pre-stimulus alpha in linguistic anticipatory processes: highly constraining linguistic contexts or cues may generate on average more precise, stronger predictions that modulate the level of pre-stimulus alpha power. Therefore, precision might be related to cue reliability in language. In case of high precision, and thus high cue reliability, when incongruent target words disconfirm the prediction a prediction error will ensue, which can be reflected in larger N400s as compared to cases in which precision is low (Bornkessel-Schlesewski & Schlesewski, 2019). Since there is evidence of associations between contextual constraint/cue reliability, pre-stimulus alpha power level, and N400 amplitude, it is possible to hypothesize that in the linguistic domain, the precision of contextual cues might affect the individuals' predictions and pre-stimulus alpha power levels, but also post-stimulus ERPs amplitude. In particular, unexpected, incongruent target words following highly precise predictions elicit a prediction error, reflected in the N400 ERP; in turn, this might be followed by an attempt to resolve the prediction error through a reanalysis process, which is instead reflected in the late posterior positivity/P600.

It should be noted that the attentional-based and precision-based interpretations of these electrophysiological dynamics are not in conflict with one another and are both compatible with a predictive account of the N400 and late posterior positivity/P600 effects. Indeed, explaining the association between pre-stimulus alpha and linguistic ERPs in terms of precision-based mechanisms incorporates and extends the attention-based interpretation, by offering a unified description of both pre and post-stimulus activity through the concept of precision.

In sum, both an attention-based account and a precision-based account point to the importance of investigating the role of spontaneous alpha oscillations in the pre-stimulus interval in relation to post-stimulus ERPs to get a glimpse of the complexity of stimulus processing dynamics. In this perspective, it is not sufficient to contrast power levels across different conditions, but it is necessary to look at variations in pre-stimulus power at the single-trial level.

1.3. The present study

The goal of the present study is to combine the investigation of traditional post-stimulus linguistic ERPs with that of pre-stimulus dynamics of alpha power. To this aim, we re-analyzed already collected EEG data (Arcara et al., 2019) on a simple semantic congruence task in which target words were congruent with the cue 50% of the time. We aimed at investigating whether trial-to-trial fluctuations in pre-stimulus alpha power can modulate the EEG response amplitude in a wide time window after each target. We did not a priori define time windows of interest so that we could detect also post-N400 effects in the most data-driven way possible, unlike previous studies on pre-stimulus alpha in language.

We hypothesized that even in absence of direct manipulations (e.g., via specific experimental design or non-invasive stimulation), different levels of pre-stimulus alpha power would modulate the N400 and P600 components (or in general any post-N400 positivity). In particular, based on the literature, for lower levels of alpha, we expected a larger N400, but we did not have any specific hypothesis for late positivity/P600 effects.

In terms of attentional mechanisms, pre-stimulus alpha levels would correspond to different degrees of anticipatory functional activation/attention (relatively low power) and inhibition (relatively high power; Klimesch et al., 2007), and therefore have an influence on how linguistic stimuli, in particular linguistic violations, are processed by individuals: a state of activation would boost the detection of mismatches between predictions and targets, producing larger ERPs; conversely, a state of relative inhibition would hinder the detection of mismatches and produce smaller ERPs.

Such a hypothesis can also be seen under a precision-based mechanism: in this case, attention is an emergent property of the precision weighting mechanism (Walsh et al., 2020) and depends on the precision of the prior (Hohwy, 2012). There is also evidence that lower pre-stimulus alpha power might reflect higher precision levels of expectations about the upcoming stimuli. Notably, as in Feldman and Friston's Posner paradigm simulation (2010), in our congruence task all cues were equally reliable; however, reliable cues can be invalid (i.e., disconfirmed by the target word). Therefore, reliability and validity are not the same thing. In the present task, the cues were equally reliable because they clearly indicated the candidate target (they presumably generated highly precise predictions). In this way, the cues varied systematically in terms of validity (they were valid 50% of the times), while their reliability/precision was theoretically constant throughout the task (cues always represented clearly a candidate target), but as stated in the previous section, the cues' characteristics do not map linearly (i.e., systematically) with the predictions' precision about the target, so that random variations of precision can take place at any level of the hierarchical inference. We assume that in this task, spontaneous alpha power fluctuations represent random variations in the ability to precisely predict the upcoming target word (Cao et al., 2017). According to the existing literature (Bauer et al., 2014; Cao et al., 2017), lower pre-stimulus alpha power might reflect higher precision levels of expectations about the upcoming stimuli, and in this condition, the mismatches between the internal model's predictions and the incongruent stimuli will result in the enhanced discrimination between expected and unexpected stimuli, therefore in higher prediction errors (Jiang et al., 2013) and greater attempts to resolve them, reflected in larger N400s and possibly, also of later components, i.e., late posterior positivities/P600.

As mentioned in section 1.1., previous studies that attempted to consider post-stimulus ERPs in relation to pre-stimulus dynamics contrasted the N400 amplitude after congruent and incongruent words within high and low pre-stimulus power conditions (or contextual constraint), thereby losing important information at the single-trial level (Rommers et al., 2017; Terporten et al., 2019; Wang et al., 2018). Indeed, based on this evidence we can only conclude that one condition is associated, *on average,* with lower pre-stimulus power and also with higher amplitude N400s, but not whether variations in pre-stimulus alpha modulate the subsequent ERPs in a trial-by-trial way. To provide an answer to this question, an appropriate statistical approach is needed. Grouping the trials within-participant and dividing them according to pre-stimulus power could be a potential solution, but it would imply converting a continuous variable (i.e., pre-stimulus alpha power) to a factorial one, e.g. by arbitrarily sorting pre-stimulus power as “high” and “low”. This dichotomous variable should then be entered as a predictor in the statistical model, but this procedure would determine the loss of important information (Cohen, 1983; MacCallum et al., 2002; Tremblay & Newman, 2015). Given the nature of the EEG signal and the goal of the study, we identified in Generalized Additive Mixed Modeling (GAMM) a method that is better suited to properly tackle the issues described above.

GAMM is a non-linear mixed-effects regression method (van Rij et al., 2019; Wood, 2016) that extends the generalized linear mixed model (GLM) by including non-linear terms, and non-linear interactions between terms, in addition to the linear ones foreseen by GLMs (Baayen et al., 2018; Baayen, Vasishth, Bates, et al., 2017; Baayen, Vasishth, Kliegl, et al., 2017; Sóskuthy, 2021; Tremblay & Newman, 2015). An important feature of GAMM is that non-linearities are modeled in a bottom-up fashion and not imposed by a priori choices. A second important aspect is that GAMMs, unlike GLMs, allow to flexibly model the interactions between continuous variables (see Baayen et al., 2010, Sulpizio et al., 2022); this is fundamental, given the interest of the present study: for each trial, we want to model how pre-stimulus power (a single value for each participant, from one continuous variable), may be associated with ERP responses, which is a time-series (a vector of values for each participant, associated with two continuous variables, Time and EEG Amplitude). A final important advantage is that GAMMs also allow to model temporal dependencies in the data, an important aspect to take into account for correct statistical inferences, and very important for EEG time-series in which each time point in the recording is highly dependent on the previous ones.

Taken together, these special features of GAMMs allow an improved fit to experimental EEG data and a systematic description of the whole time window of interest, rather than only focusing on the averages in predefined time windows, or the statistical significance of the differences between the experimental conditions (van Rij et al., 2019). This aspect is essential in the present study, which aims at analyzing the effects of pre-stimulus alpha power in a wide time window starting from the target word onset, including both early and late language-related ERP components, without a-priori defining them. Because of their flexibility, GAMMs have already been used in the analysis of EEG data in linguistic experiments (e.g., Alday et al., 2017; De Cat et al., 2015; Meulman et al., 2015; Porretta et al., 2017; Tremblay et al., 2016; Tsiwah et al., 2021).

2. Materials and Methods

2.1. Participants

27 participants volunteered to take part in the study, after signing an informed consent form. One participant was excluded from the analysis because of a misunderstanding of the task instructions and 5 participants did not complete part of the recording (see section 2.2.). Thus, the final analyses were performed on 21 participants (15 females; mean age 25.1 years). All participants were right-handed, native speakers of Italian, had normal or corrected-to-normal vision, and reported no history of reading or learning disorders. The experiment was approved by the Local Ethics Committee. Data presented in this study comes from a larger experiment aimed at investigating agreement violations in number morphology, with already published results (Arcara et al., 2019). The published results did not affect in any way neither the hypotheses of the present study, nor the statistical choices

made. Here we will only cover the parts and materials of the experiment that are relevant to the present study and that were used in the present analyses.

2.2. Procedures

Participants were tested in a dimly lit, quiet room. They were asked to complete a picture-phrase matching task, performed on a computer screen. The task was delivered with the E-prime software (Psychology Software Tools, 1999, Pittsburgh, PA). Each trial was structured in the following way: first, a fixation cross appeared in the center of the screen (1000 ms); afterward, a picture (the cue) showed up (1000 ms), followed by a short blank screen for 200 ms and then by two Italian words. The first word (the quantifier) was displayed for 300 ms, followed by a blank screen (200 ms), and the second word (the target) was displayed for 300 msec. The quantifier was always congruent with the cue, while the target could be congruent or incongruent with the cue (with a 50% proportion). The words were followed by another blank screen with a random duration between 1000 and 1500 ms, after which two response words (True and False) appeared on the right and left sides of the screen. Participants were asked to respond whether the two-word sequence was congruent or incongruent with the preceding picture, without any time pressure. The position of the response words (i.e., True/False), as well as that of the response keys, were counterbalanced across participants. The trial procedure is illustrated in Fig. 1. All stimuli subtended at most 5 degrees on the horizontal plane to avoid excessive eye movements. miniFive practice trials were administered before the beginning of the experiment to familiarise with the task. The overall task lasted about 45 minutes and included twelve breaks. Before the beginning of the task, a 5-minute session of an eyes-open resting state was recorded, which was used for normalizing the pre-stimulus alpha values, as described in Section 2.5.

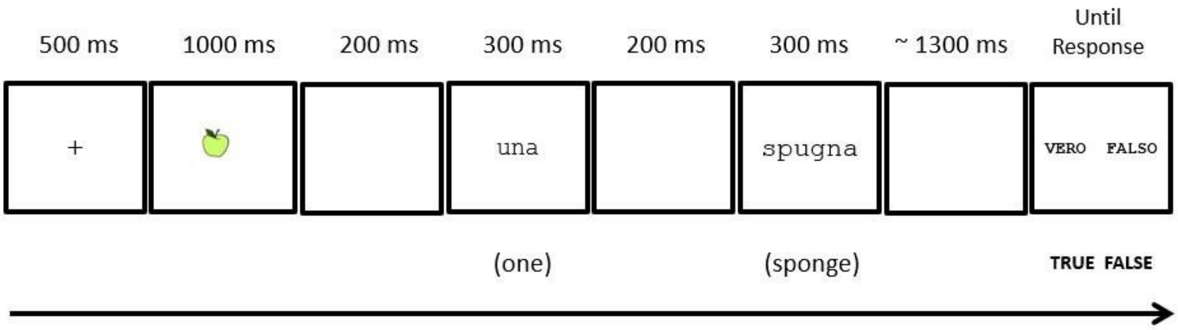


Figure 1. Task Design. All trials followed the depicted sequence: after a fixation cross, a picture was displayed, followed by a two-word phrase presented in a word-by-word fashion. In the picture, “una” = “one”, while “spugna” = “sponge”. Participants had to respond if the phrase matched the presented picture by pressing two buttons associated with TRUE/FALSE response (FALSE, in the depicted example; “vero” = “true”; “falso” = “false”). There was no time pressure for the response.

2.3. Materials

The linguistic stimuli of the experiment consisted of phrases made up of quantifier-noun pairs and three Italian quantifiers were selected: un/uno/una, alcuni, and qualche, corresponding to a/an, some+plural inflection of the noun, and some+singular inflection of the noun (which is perfectly grammatical in Italian; for the linguistic details of the stimuli, see Arcara et al., 2019). The stimuli were created to be matched as much as possible for length and frequency. In each phrase, all content nouns referred to concrete, countable, and non-animate objects. We selected two pictures for each noun, representing either one single object or four instances of that object. The drawings in the pictures were arranged to avoid any kind of effect due to structural composition. Each picture-to-phrase matching could be congruent (e.g., a picture of four apples followed by the phrase ‘some apples’) or incongruent (e.g., a picture of one orange followed by the phrase ‘one sponge’). The mismatches considered in the present analyses only concerned the referential objects, but the experiment included also other experimental conditions related to number morphology that were excluded from the present analyses (see Arcara et al., 2019, for a complete description of the task). Each combination of experimental variables included 30 stimuli for a total of 180 experimental stimuli. In the present analyses, we focused on the case in which there was either a semantic congruence or a semantic incongruence between the picture and the following words, as it is the one in which a higher (and more easily interpretable) N400 is expected.

Statistical analyses were conducted only on trials with correct behavioral responses.

2.4. EEG data recording and analysis

The EEG signal was recorded from 28 active electrodes embedded in an elastic cap, arranged according to the 10/20 system (Brain Products, Acticap). Each electrode was referenced online to the left earlobe. Three additional electrodes were used to monitor eye movements and blinks, with two electrodes placed near the outer corner of the eyes (external canthi) and one placed in a pupil-centered position, under the left eye. The impedance of each electrode was kept lower than 10 KΩ throughout the recording. The following electrodes were included: Fp1, Fp2, Fz, F3, F4, F7, F8, FC1, FC2, FC5, FC6, C3, C4, Cz, T7, T8, CP1, CP2, CP5, CP6, P3, P4, P7, P8, Pz, O1, O2, Oz. The EEG signal was amplified by using BrainAmp amplifiers with a hardware high-pass of 0.1 and with a sampling rate of 500 Hz. EEG data were pre-processed with the Brainstorm MATLAB toolbox (Tadel et al., 2011; March 2015 version). In the pre-processing phase, we applied a high-pass filter at 0.5 Hz to the continuous data and then used Independent Component Analysis (ICA) for artifact removal. From the ICA-corrected continuous data, we extracted epochs time-locked to the onset of the target word (content noun), ranging from -3020 msec to 1470 msec after the stimulus and including the quantifier (cue). Trials were baseline corrected to the mean value of 100 msec preceding the target word, and those containing excessive artifacts were rejected in this phase after visual inspection. For each condition (congruent and incongruent), 94.5% of trials were accepted on average (on average, 85.05 out of 90 trials per condition). There were no appreciable differences across conditions: the number of accepted trials separated for conditions ranged from 94% to 95%. The single-trial data were subsequently downsampled to 100 Hz to reduce computational burden and filtered with a 50 Hz low-pass to avoid aliasing. Data were thus exported to R (R Core Team, 2013) with the package erpR (Arcara and Petrova, 2017) for subsequent statistical analyses. Custom code was developed for data handling and plotting.

Resting-state preprocessing was identical to that performed for the task. After the ICA artifact removal, continuous recordings were split into 5 seconds epochs and trial rejection was performed based on visual inspection as in the task recordings.

2.5. Single-trial time-frequency analysis

After trial rejection, time-frequency (TF) analysis was performed with Brainstorm (Tadel et al., 2011, ver. September 2018) on epoched, non-downsampled data. A Morlet wavelet was constructed, with a central frequency of 1 Hz and 3 s of time resolution (full width half maximum, FWHM). The wavelet ranged from 1 to 45 Hz, with 1 Hz linear frequency steps. We chose this frequency range to enable other analyses beyond the scope of this paper, that are not going to be presented here. From the single-trial time-frequency decomposition, we then extracted the average magnitude in the alpha band (8-13 Hz) and exported the single-trial alpha magnitudes to R for statistical analyses. The same procedure was applied to the resting-state recordings (previously epoched in 5 seconds segments for trial rejection purposes), with the exception that time-resolved alpha-band magnitudes in the single epochs were first averaged in time (to obtain a Morlet-based frequency spectrum), then averaged across epochs and then exported. We chose to use the average resting alpha power as a baseline, instead of other inter-trial intervals, because subjects were instructed to blink during such intervals (to reduce contamination during the experimental trials). As a consequence, power estimates from this time window would have been influenced by such artifacts.

We obtained pre-stimulus power by averaging the TF values across the time window between -200 and -10 msec before the target word. We chose this interval, albeit short, because it was the closest event-free interval to the target word, therefore enabling us to investigate pre-stimulus dynamics. Ending the time-window 10 ms before the target allows to avoid any undesired effect of temporal smearing due to the time-frequency analyses, i.e., the temporal leakage of stimulus-related oscillatory activity to the pre-stimulus time-window, in particular if the stimulus occurs at time 0 or shortly thereafter (Cohen, 2014).

We thus obtained a single mean value of pre-stimulus alpha power for every electrode and trial. As the absolute power value could be related to irrelevant factors (e.g., individual characteristics of the participants, quality of the signal, etc.), we normalized pre-stimulus power with a baseline division procedure, consisting of dividing power during the task by the average power during the baseline period (in our case, resting-state activity). The resulting unit of this procedure is a ratio (Cohen, 2014). Values smaller than 1 mean that in each pre-stimulus period, a given subject had a lower alpha power than during resting; values larger than 1 mean that the subject's pre-stimulus power was higher than the resting state power. Since the power distributions were skewed, we log-transformed the power values for statistical analysis purposes.

2.5.1. Individual Alpha Peak Frequency

Characteristics of alpha oscillations (e.g., peak and power) vary across individuals (Wang et al., 2022), so researchers have recently directed their attention to a particular feature of the alpha band signal, the individual alpha peak frequency (IAPF), defined as the peak frequency in the alpha range with the greatest power (Bazanova & Vernon, 2014). However, the previous studies on pre-stimulus alpha power and N400 have been conducted considering the traditional 8-13 Hz alpha band, so we ran the same GAMM model on two versions of the same dataset, one obtained with the standard alpha band to be consistent with the existing literature, and one with the IAPF.

We obtained the IAPF from the resting state data by first computing the power spectrum densities (PSDs) on the continuous signal. PSDs were subsequently fed to the Brainstorm function *specparam*, which estimates and subtracts the 1/*f* noise from the PSDs and allowed us to extract the peak frequency within 8 and 13 Hz for each subject and electrode. In case an electrode had two peak frequencies in a given subject, we chose the peak frequency with the higher power (Katyal et al., 2019). On the other hand, some electrodes may lack a peak in the alpha band. To solve this issue, we calculated the 75% quantile of the number of electrodes (n=21) and adopted this value as a threshold, so that all subjects with less than 21 electrodes with a peak in the alpha range were excluded from further analyses (n=4). As a result, IAPF analyses were performed on 17 subjects. For each subject, the global IAPF value was obtained as the median of all electrodes' IAPF values. Pre-stimulus individualized alpha power was then calculated on the task data at the single trial level within ±1.5 Hz around this IAPF value, with the same procedure described in Section 2.5.

2.6. Statistical analyses

The analyses were performed with R (version 3.4.4) and were focused on the ERPs time-locked to the target word (the content noun, see Figure 1).

2.6.1. Preliminary analysis.

We obtained averaged ERPs for conditions, by averaging all trials within participants, and then averaging data from participants (see Figure 2). We conducted some preliminary analyses to investigate whether potential and spurious differences in pre-stimulus alpha across conditions were present and that could explain the difference between congruent and incongruent trials (we do not expect to find such differences, because the congruence of a trial was defined only by the target word). To this aim, we compared the average pre-stimulus alpha power ratio (see section 2.5.) across congruent and incongruent conditions for each electrode with a within-participants t-test, corrected for multiple comparisons with the Bonferroni method.

2.6.2. Generalized Additive Mixed Models (GAMMs)

After exporting the single trials ERPs to R, we restricted the time window of interest from -100 to 800 msec. Data from each electrode were analyzed separately (as in de Cat et al., 2015), using the same model syntax, in which continuous ERP amplitude in the whole time window of interest was entered as the dependent variable (before the model described here, several other options were run and compared). Main effects included:

- condition as a linear predictor;
- a non-linear effect of time, depending on condition (this term captures the different changes in the ERP over time in the two conditions, i.e., the N400 and P600 effect);
- a non-linear effect of pre-stimulus alpha power, capturing the (possibly) non-linear modulation of ERP amplitudes by different magnitudes of pre-stimulus alpha.

An interaction term, specifying the non-linear interaction of interest between time and pre-stimulus alpha power depending on the condition, was included to capture whether pre-stimulus power modulates, in a possibly nonlinear way, the subsequent ERP amplitudes, in either of the two conditions. In GAMMs, interactions between continuous variables are modeled by tensor smooth functions.

To reduce autocorrelation in the model's residuals and account for dependency in the data, we added a random structure as recommended by Van Rij and colleagues (2019). The random structure should include random factor smooths for participants and items and a random intercept and slope for each time series (i.e., item). Such structure essentially adjusts the model's predictions for every single trial of each subject. However, doing so would result in too large computational demands; in addition, during the EEG recording, we did not keep track of which item (i.e., target word) was presented in each trial. We, therefore, constructed the random structure attempting to reach a trade-off between the model's accuracy and the computational load. To this aim, we included two random factor smooths: one for participants and one for each combination of quantifier and condition, instead of each item. We then created a factor variable, called *Event*, that identified each trial through the unique combination of participant, quantifier, and trial number. We finally included this variable as both a random intercept and slope (in interaction with time). Please note that the quantifier was only included in the random structure to model part of the random variability between trials, and is not a variable of interest in the analysis; therefore, it will not be discussed further.

After specifying the model syntax, we tested the model on a randomly selected electrode. Unsurprisingly, the residuals showed some remaining autocorrelation. This could be related to the filters applied to the signal: as filtering can be conceived as a weighted average of the signal at subsequent time points (Luck, 2014; for detailed considerations on downsampling and filtering in GAMM, see Van Rij et al., 2019). We dealt with this issue by including an AR1 model that accounts for autoregressive processes in the data. We estimated the value of *rho* as the model's autocorrelation function (ACF) at Lag 1 and we included it in our model, together with a term specifying the starting point of each time series. The models' nonlinear random structure and the AR1 error model are crucial for dealing with the intrinsic temporal dependency in time series data (van Rij et al., 2019). As the model's residuals were not normally distributed, we fitted the model with a link function for a scaled-t distribution (Wieling, 2018). Since adding this function considerably increases computational time, we set the argument "*discrete*" to true, enabling more efficient processing.

The final model R syntax is reported here:

```
Ampl ~ Cond + s(Time, by=Cond) + s(norm_log_power)  # main effects

+ ti(Time, norm_log_power, k = c(40, 40), by = Cond)  # tensor with
main interaction

+ s(Time, ID, bs="fs", m=1) # nonlinear random smoother with
individual differences in time course
```

```
+ s(Time, cond, bs="fs", m=1)   # nonlinear random smoother with
differences for original condition (should have been item)
+ s(Event, bs="re")  # random intercept for individual time series
+ s(Time, Event, bs="re") # random slope for individual time series
```

The same model was run on the IAF dataset, only adjusting the number of basis functions to fit the lower number of subjects included.

2.6.3. Visualization of results

In GAMMs, the effects of the interactions of interest are mainly interpreted through visual inspection of the tensor functions' plots, which can be 2- or 3-dimensional, depending on the number of numerical predictors (de Cat et al., 2015). In this case, our main effect of time depending on condition (the term "s(Time, by=Cond)") has one numerical predictor (time), which captures the different ERP waveforms (i.e., voltage amplitude) of congruent and incongruent trials and results in a 2-D plot with time on the x-axis and amplitude on the y-axis (Figure 2), that allows to link GAMM results to the typical waveforms studied in the ERP literature (Sulpizio et al., 2022). On the other hand, the main interaction term "ti(Time, norm_log_power, k = c(40, 40), by = Cond)" has two numerical predictors (time and pre-stimulus alpha). Therefore, it results in 3-D plots with time on the x-axis, pre-stimulus alpha power level on the y-axis, and color-coded voltage amplitude (Supplementary Figures 1S and 2S) that provide information on how ERP waveforms are modulated by pre-stimulus power depending on congruent and incongruent conditions. This interaction allows us to visualize the voltage amplitude depending on pre-stimulus alpha levels and conditions, but also to perform subtractions between the incongruent and congruent tensor surfaces as if they were ERP waveforms. We followed this procedure because comparing experimental conditions with their raw ERPs is not an optimal strategy, and difference waves allow us to overcome this difficulty by isolating components of interest that are informative on the effects of the experimental manipulations (Luck, 2005). Moreover, subtracting the incongruent from the congruent tensor surface also allows the identification of significant N400 and other effects (predicted differences between voltage amplitudes) in relation to the pre-stimulus alpha level. These effects were calculated and plotted with the plot_diff2 function in the package itsadug (Van Rij et al., 2020) and are shown in Figure 4 and the upper panels of Figure 5A-D. Since the interpretation of 3-D difference plots can be confusing, it is possible to break them down by selecting different values of a variable (in our case, pre-stimulus alpha power) and plot 2-D graphs "zooming in" on the interaction of the remaining variables (in our case, time and voltage amplitude), highlighting the differences between conditions for different values of pre-stimulus alpha. We show this in the lower left panels of Figures 5A-D. The lower right panels, instead, show the time windows where the differences between conditions for different values of pre-stimulus alpha are significant.

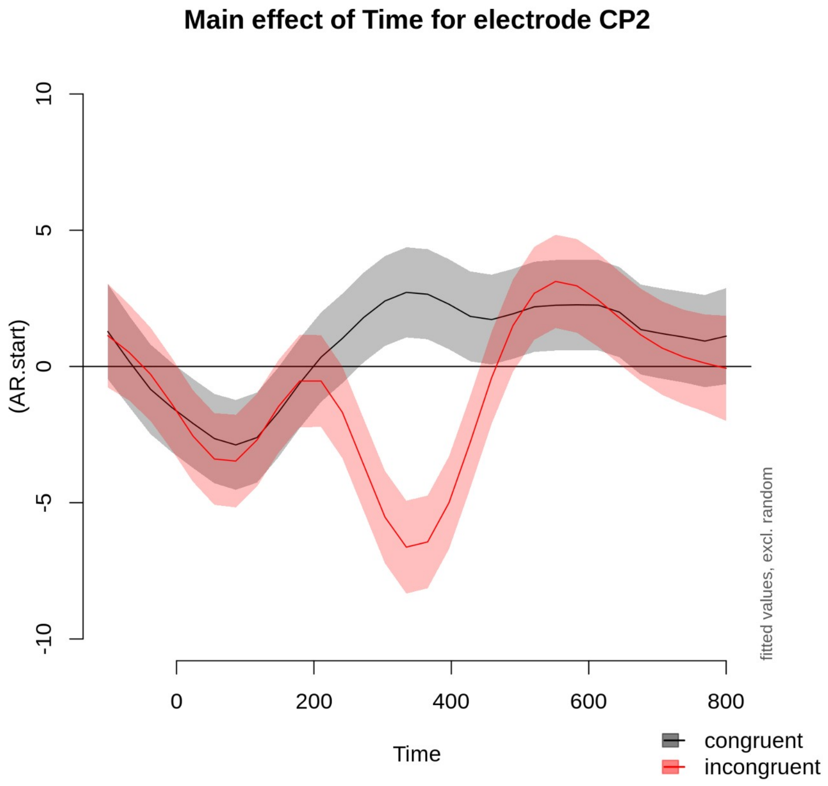

Main effect of Time for electrode CP2

Figure 2. Main effect of Time for one exemplificative electrode (Cp2). The term "s(Time, by=Cond)" captures the different ERP waveforms (i.e., voltage amplitude) of congruent and incongruent trials and results in a 2-D plot representing the congruent and incongruent tensors, which allows to link GAMM

results to the typical waveforms studied in the ERP literature. 0 ms in the X axis (Time) represents the target stimulus onset.

To explore whether a more traditional analysis would have captured the modulations associated with pre-stimulus alpha we performed some additional analyses using ANOVAs instead of GAMM. These analysis and their results are reported in the Supplementary materials. The code used for the analys can be found at the following OSF link: https://osf.io/6jrk2/?view_only=00d6b579199a47758d8b8a2b854dabcf

3. Results

3.1 Behavioral results

As in the original paper (Arcara et al., 2019), the performance in the task was almost at ceiling in all subjects. The mean percentage of errors was .8% (mean number or errors = 2.96, SD = 3.513, range = 0-13). As the performance was at ceiling, data on accuracy were not further analysed. As there was no time pressure to give the response, reaction times were not analysed.

3.2 ERP results

ERP grand averages across conditions is shown in Figure 3. The similarities between the grand averages and the tensors (main effects) of Time depending on conditions (Figure 2) indicate that the statistical model provided a good fit to the data. Preliminary analyses investigating the potential presence of spurious results did not show any significant effect, indicating that congruent and incongruent trials did not differ for pre-stimulus alpha power.

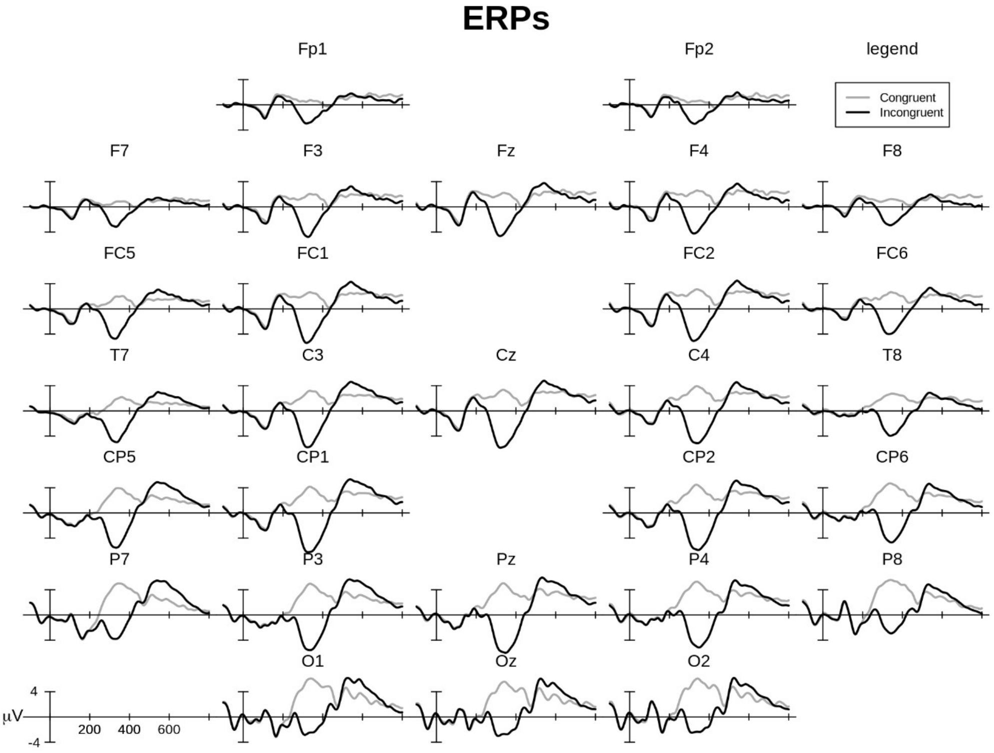


Figure 3. ERPs grand averages for all electrodes and conditions.

Table 1 reports the significance of the interaction between time, power, and congruent vs incongruent conditions for each electrode, in the model using the standard alpha band (8-13 Hz). Analyses on IAPF yielded very similar results hence only results on the standard alpha band are reported here (see Supplementary materials for details).

The percentage of deviance explained by the models ranged between 16.3 and 24.1. The interaction between time and pre-stimulus alpha power for the incongruent condition was significant for almost all electrodes (Fp1, Fp2, F3, T7, C4, CP5, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2, Oz). The same interaction for the congruent condition was significant in almost all electrodes (Fp2, F3, Fz, F8, FC5, FC1, C3, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2, Oz). This means that pre-stimulus alpha power significantly predicted the ERPs modulation in both the congruent and incongruent conditions in almost all electrodes, especially centroparietal and posterior ones (see Figures 1S and 2S).

Table 1. Significance of interactions between time, power, and condition for the GAMM on the standard alpha band.

Electrodes	Interaction: time, power, incongruent condition	Interaction: time, power, congruent condition
Fp1	0.02*	0.12
Fp2	0.01*	0.05*
F7	0.77	0.70
F3	0.03*	< 0.001*
Fz	0.20	< 0.001*
F4	0.53	0.22
F8	0.31	0.03*
FC5	0.08	0.00*
FC1	0.07	< 0.001*
FC2	0.30	0.20
FC6	0.40	0.08
T7	0.05*	0.09
C3	0.13	0.01*
Cz	0.24	0.18
C4	< 0.001*	0.02*
T8	0.07	0.14
CP5	< 0.001*	0.03*
CP1	0.45	0.02*
CP2	< 0.001*	0.04*
CP6	< 0.001*	0.01*
P7	< 0.001*	< 0.001*
P3	< 0.001*	< 0.001*
Pz	0.01*	0.01*
P4	< 0.001*	0.02*

P8	< 0.001*	< 0.001*
O1	< 0.001*	< 0.001*
Oz	< 0.001*	< 0.001*
O2	< 0.001*	< 0.001*

Significant results are marked with *.

Figure 4 depicts, for each electrode, the predicted amplitude (tensor surface) difference between incongruent and congruent ERPs, i.e., the N400 and P600 effects depending on the pre-stimulus alpha level. Time is represented on the x-axis, pre-stimulus alpha power level on the y-axis, and predicted voltage amplitude difference is color-coded. Colored blots represent significant differences, while white areas indicate regions where the confidence intervals (95% CI) around the predicted surface included zero. Shades of blue and green represent negative differences, while shades of red and yellow stand for positive differences.

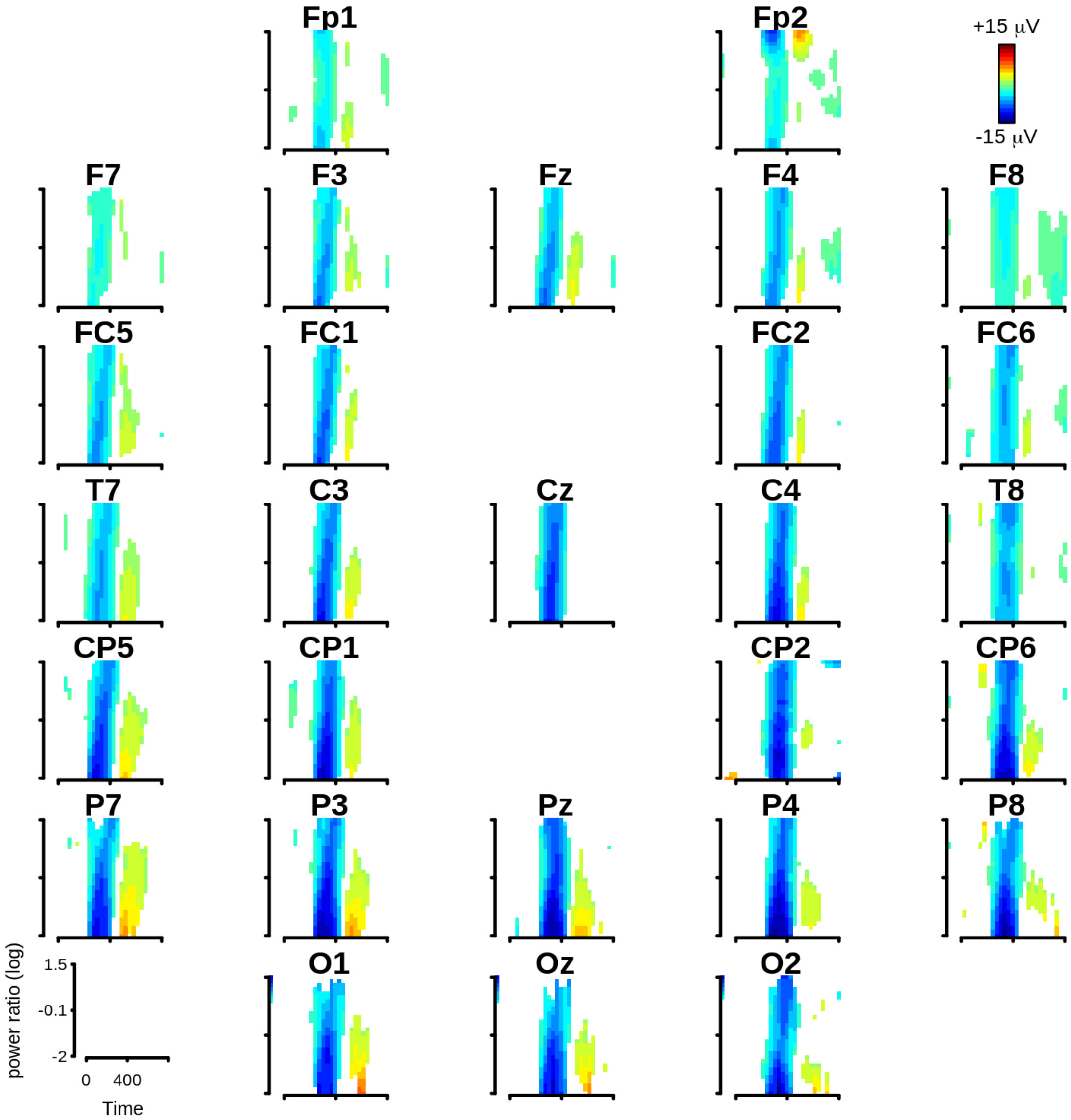


Figure 4. Predicted amplitude (tensor surface) differences between incongruent and congruent conditions for the main interaction of interest (time, alpha band power, and conditions). The figure depicts the N400 and P600 effects depending on the pre-stimulus alpha level. Time is represented on the x-axis and pre-stimulus power on the y-axis. Colored areas reflect significant predicted ERP amplitude differences between congruent and incongruent responses. Shades of yellow and red indicate increasingly positive differences, while shades of green and blue indicate increasingly negative differences. Green shades indicate that the difference tends to zero. White areas indicate regions where the confidence intervals (95%) around the predicted surface included zero.

Power modulated the differential surfaces in slightly different ways across scalp locations, but the predicted amplitude difference between congruent and incongruent ERPs was significant in time windows compatible with those of the N400 and a later positive component resembling a late posterior positivity/P600. This effect was observable for almost all electrodes, as shown in Figure 4 by the colored blots indicating that the 95% CI around the predicted amplitude difference did not include zero. The effect was less marked at frontal rather than centro-parietal sites. However, the predicted amplitude in the N400 time window appeared to be significantly modulated by pre-stimulus power, especially over central and posterior electrodes. The negative effect in the N400 time window was observable in all electrodes regardless of power, but its amplitude was more negative for lower pre-stimulus power (darker blue shades).

At all locations except Fp2 and Cz, the predicted amplitude in the late posterior positivity/P600 time window was significantly more positive for lower power and was maximal over left centro-parietal electrodes (yellow shades); frontally, it was followed by weakly negative differences, differently modulated across electrodes. A brief negative deflection appearing immediately post-stimulus was visible in electrodes Fp1, F7, and T7, where it was associated with lower power; contrarily, in CP1 the opposite pattern was visible.

To better characterize the effects, Figure 5A-D shows the highlight for some specific electrodes.

The top panels in Figure 5A-D show the tensor surface difference between responses to congruent and incongruent targets, as in Figure 4. To allow a more detailed inspection of such difference, we broke it down by selecting three different values of pre-stimulus alpha power (high, median, and low, corresponding to the upper quartile, the median value, and the lower quartile, respectively) and plotted accordingly the smooth functions for congruent and incongruent conditions (bottom left panels), as well as the difference between congruent and incongruent smooths and the time windows where it is significant (bottom right panels).

Overall, for low and median power, the predicted amplitude difference was significant for longer time windows than for large power, especially for the time window corresponding to the late posterior positivity/P600 effect (see electrodes C3, C4, P3, P4). Sometimes significant differences in this time window were only present for low power (see electrode CP2). For CP1, we can also observe significant positive differences immediately post-stimulus, for both median and high power.

In summary, results with GAMM showed that pre-stimulus alpha power significantly modulated the amplitude of the N400 and late posterior positivity/P600 effects. In particular, the N400 effect was present regardless of the pre-stimulus alpha level but was more negative for lower levels of pre-stimulus alpha. The P600 effect, instead, was only significant for lower levels of pre-stimulus alpha power. Overall, the pattern of significant differences resembles the ERPs' grand average in Figure 2.

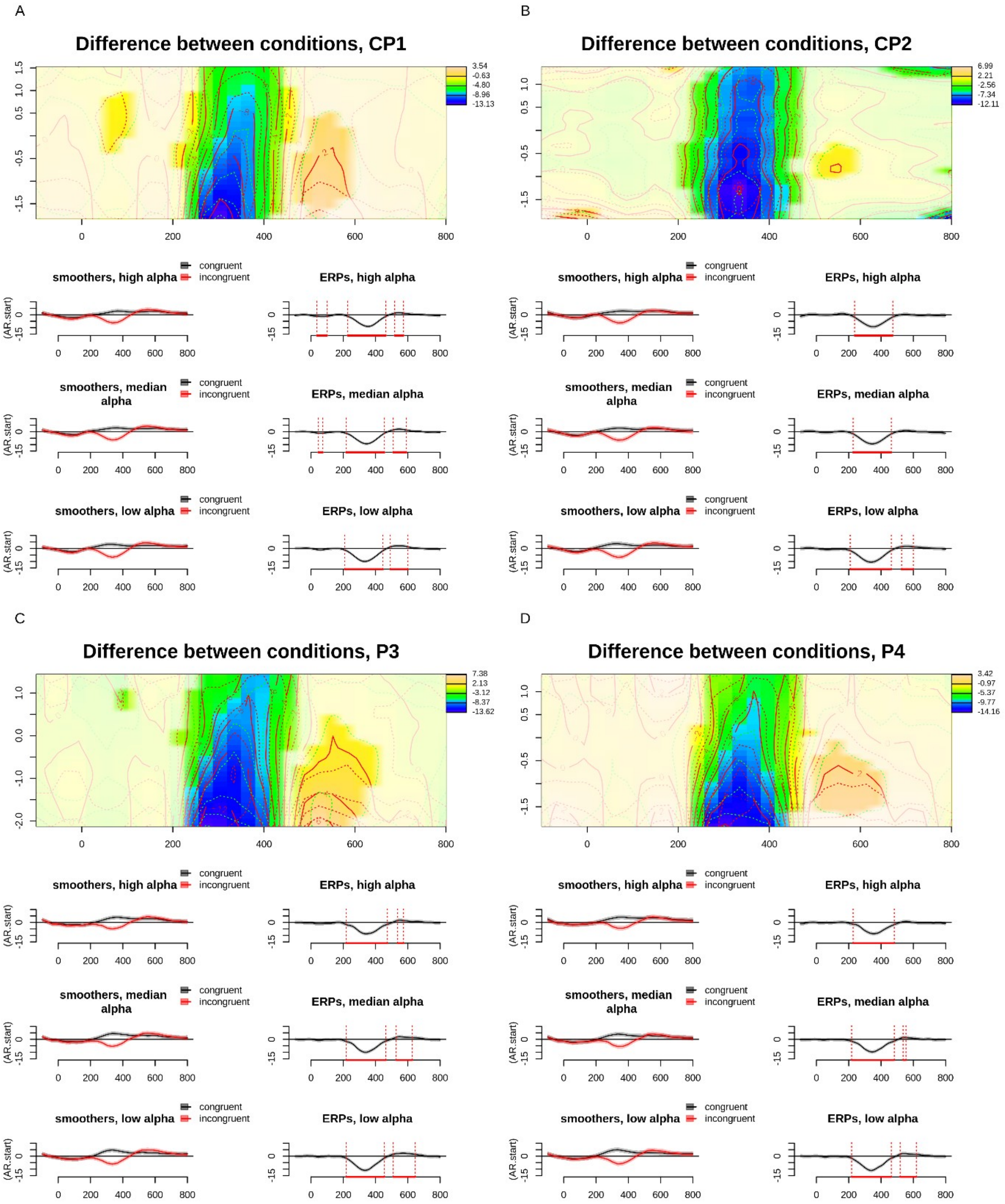


Figure 5. Highlights for specific electrodes. White areas indicate regions where the confidence intervals (95%) around the predicted surface included zero.

4. Discussion

In the present study, we investigated whether and how trial-to-trial spontaneous fluctuations in pre-stimulus alpha power modulate the EEG response amplitude in a wide time window after congruent and incongruent target words, during a semantic congruence task. Compared to the previous literature, we did not experimentally manipulate the level of pre-stimulus alpha power, but we focused on its spontaneous fluctuations, which could be interpreted according to an attention-based mechanism (cortical activation/inhibition; Klimesch, Sauseng, & Hanslmayr, 2007; Foxe and Snyder, 2011; Jensen & Mazaheri, 2010) or according to a precision-based mechanism (Bauer et al., 2014; Cao et al., 2017; Walsh et al., 2020). The latter posits that the stimulus' precision levels could be linked to cue reliability in language, so that when predictions are disconfirmed, the prediction error would be manifested as more negative N400s (Bornkessel-Schlesewsky and Schlesewski, 2019). Focussing on the trial-to-trial alpha fluctuations has allowed us to perform analyses at the single-trial level, thereby avoiding the contrast between two conditions, that would have given results relative in nature and entailed the loss of important information (Cohen, 1983; MacCallum et al., 2002; Tremblay & Newman, 2015; Van Diepen et al., 2019). Differently from previous studies, we also avoided the a priori definition of time windows of interest, so that we could explore the possible influence of pre-stimulus alpha also at time points that had not been considered in previous literature (e.g., post-N400 positivities), in a data-driven way. To these aims, we used Generalized Additive Models, that allow to model complex non linear relationships in a mixed model framework. A control analysis using a more traditional ANOVA approach did not yield any effect, supporting the importance of using GAMMs for the analysis (see also Supplementary Materials, p. 19).

Results of the study suggest the presence of an N400 effect with a typical centro-parietal distribution and a later positive component also centro-parietally distributed, which resembles a late posterior positivity/P600 (Figure 3). It is important to point out that late posterior positivities/P600 are mostly found in sentence reading paradigms, where participants read structured sentences and have to integrate anomalous targets into rich linguistic contexts (Brothers et al., 2020), which is not the case for this experiment. However, there is also evidence that these potentials are associated with the detection and reanalysis of linguistic anomalies in the broad sense (see Section 1.1.; Quante et al., 2018; Van Petten & Luka, 2012) and according to predictive processing theories, they follow prediction errors to resolve them. According to this definition, these processes could also happen in more basic linguistic tasks other than sentence reading, so we included in our hypothesis also post-N400 positivities and, in particular, the late posterior positivity/P600. As the experiment design is not the one that would be typically used to elicit a classical P600 (see also Arcara et al., 2019), it is difficult to clearly interpret the functional meaning of the effects found in a late time window. Nonetheless, the effect was present and compatible with accounts of late posterior positivities/P600 as correlates of detection and reanalysis of linguistic anomalies; therefore, for clarity reasons, we will refer to our late positive component as a classic late posterior positivity/P600 mirroring these processes, while in a predictive processing framework, it could reflect the attempt at resolving a prediction error, which, we speculate, would not be strictly related to the type of task, but to the incongruence itself. Our results also point to a complex dynamic showing that the N400 and the late posterior positivity/P600 found in centro-parietal electrodes were larger in the case of lower pre-stimulus alpha power. However, the N400 was observable regardless of pre-stimulus power level, albeit more negative for lower pre-stimulus power values (as shown by the comparison of tensor surfaces, see Figure 4 and upper panels of Figure 5A-D), while the late posterior positivity/P600 effect was only observable for low pre-stimulus alpha. These findings extend previous reports on the influence of pre-stimulus alpha power on the linguistic ERPs and demonstrate that the pre-stimulus state exerts a stronger modulation on later components such as the P600 than on earlier ones, such as the N400.

We also replicated the analyses using the IAPF power, which yielded similar results to those obtained with standard alpha band power (see Supplementary Materials, p. 216 and following). In the language domain, it seems that overall, a lower IAPF is associated with more efficient linguistic processing (e.g., Bornkessel-Schlesewsky et al., 2022; Nalaye et al., 2022; for further considerations, see the Supplementary Materials, p. 15). This suggests that properties of alpha oscillations other than power, such as IAPF speed, might be predictive of post-stimulus cognitive processes and electrophysiological responses. This field therefore deserves further investigation.

In sum, the results confirm our initial hypothesis, positing lower pre-stimulus power associated with larger N400s and late posterior positivities/P600. In addition, they showed that pre-stimulus alpha power differently modulated the N400 and late posterior positivity/P600 effects in the semantic congruence task. Such differential modulation could be rooted in the different functional roles of the ERPs.

It has been recently hypothesized that the N400 and the late posterior positivity/P600 could share some similarities with ERPs that are classically found in perceptual tasks, i.e., the mismatch negativity (MMN) and P300 respectively and that their increased latencies and different topographies reflect the increased complexity of linguistic stimuli when compared to perceptual ones (Bornkessel-Schlesewski and Schlesewski, 2019; Leckey & Federmeier, 2020). In the following paragraphs, we will discuss the possible interpretations of our results in line with the cognitive processes reflected in the N400 and the late posterior positivity/P600, under both an attention-based mechanism and a precision-based one.

4.1. Cortical inhibition and attention

In response to perceptual stimuli, the MMN is associated with the degree to which the encountered stimulus is incongruent with (or deviates from) the previous ones and is elicited even when participants do not pay attention to the stimuli (Garrido et al., 2009). In the present study the N400 is present regardless of the pre-stimulus alpha power level, but is more negative for lower levels of pre-stimulus power, so we can argue that this ERP might reflect processes that are relatively independent of the pre-stimulus state, interpreted as a correlate of the level of attention or cortical activation in the pre-stimulus interval. Assuming a similarity between the N400 and the MMN, it is possible to hypothesize that the N400 reflects automatic ways to signal incongruent and salient linguistic stimuli, and is only partially modulated by the pre-stimulus state (attention or level of cortical activation).

The late posterior positivity/P600, instead, deserves separate consideration. In line with the analogy between perceptual and linguistic ERPs, in the perceptual domain the P300 is related to attentional reorienting following infrequent, salient, and task-relevant stimuli (Corbetta, Patel, and Schulman, 2008; Polich, 2007; Sassenhagen & Bornkessel-Schlesewsky, 2015). In this domain, the P300 is also modulated by pre-stimulus alpha power: lower pre-stimulus power is linked to larger P300 amplitudes (Ergenoglu et al., 2004), while higher pre-stimulus power to longer latencies and attenuated amplitude (Grent-'T-Jong et al., 2011; Price, 1997; see Section 1.2.). During language processing, linguistic anomalies such as semantic incongruencies are usually infrequent, salient, and task-relevant. Therefore, they share those characteristics of the stimuli that, in attention/perception domains, elicit the P300; also, this ERP and the late posterior positivity/P600 may reflect, at least in part, the same neurocognitive processes (e.g., response to salience; Sassenhagen & Fiebach, 2019). In this vein, our results show that also the late posterior positivity/P600 amplitude is larger for lower pre-stimulus alpha power, thus substantiating hypotheses about the similarity between the P600 and P300. In those studies linking pre-stimulus alpha and the P300, lower alpha power was interpreted as a correlate of attention or cortical activation, so we can argue that also in the present study anticipatory attention/cortical activation possibly have an influence over the P600 and the underlying cognitive process in response to semantic incongruencies. In addition, semantic incongruencies also trigger the reanalysis of linguistic stimuli, in an attempt to make sense of them; as we mentioned in Section 1.1., the late posterior positivity/P600 could precisely reflect such a process. A possible explanation for the influence of the pre-stimulus state on linguistic reanalysis would be that this process might entail the allocation of attentional resources to some extent; therefore, it could be dependent on the level of cortical activation or attentional resources that were induced by the cue in the pre-stimulus interval, as indicated by the level of pre-stimulus alpha power: the lower the power, the greater the attentional focus, which results in a larger post-stimulus late positive positivity/P600. In other words, the amplitude of this positivity varies spontaneously as a function of the level of pre-stimulus alpha power, which may reflect the level of anticipatory attention/cortical activation deployed to stimulus analysis before the presentation of the incongruency.

4.2. Precision and attention

The analogy between perceptual and linguistic cognitive domains and ERPs also suits a more parsimonious predictive processing perspective, where the perceptual MMN is thought to be an early signature of sensory prediction errors, while the later P300 is elicited in response to violations of higher-level regularities (Garrido et al., 2021). Similarly, the N400 mirrors the degree to which the actual linguistic stimulus deviates from the individual’s expectations, i.e., the semantic prediction error, which essentially signals a higher-order violation that needs to be resolved through a reanalysis process, reflected in the late posterior positivity/P600 (Kuperberg et al., 2020; Wang et al., 2021). Tellingly, predictive processing accounts consider N400s and late posterior positivities/P600s as part of the same process of prediction error generation and resolution, through the concept of precision. Precision, the inverse of variance or level of reliability of a hypothesis, is related to cue reliability in language (Bornkessel-Schlesewski & Schlesewski, 2019). However, precision levels fluctuate spontaneously during the task even in presence of equally reliable cues (Feldman and Friston, 2010) and there is evidence that higher precision levels are associated with lower pre-stimulus power (Bauer et al., 2014; Cao et al., 2017). In section 1.1 we reviewed evidence stating that in sentence comprehension paradigms featuring different levels of contextual constraint, more constraining contexts were associated with lower levels of pre-stimulus alpha power (Gastaldon et al., 2020; León-Cabrera et al., 2022; Rommers et al., 2017; Terporten et al., 2019; Wang et al., 2018), while similar studies found that contexts eliciting lower pre-stimulus alpha power were associated with more negative N400s (Gastaldon et al., 2023; Rommers et al., 2017; Terporten et al., 2019; Wang et al., 2018). This body of evidence helped us generate the hypothesis that lower pre-stimulus alpha might be associated with more precise predictions. However, the paradigm we used in the present study presents some important differences from sentence comprehension paradigms. In particular, we did not manipulate the level of pre-stimulus alpha power or prior precision: the graphic cues were equal in terms of reliability, so participants did not need to generate different predictions for lexical-semantic candidate targets suggested by a sentence context, as the candidate target was clearly cued by the picture (e.g., picture of an apple). Thus, in all trials, participants could pre-activate the specific lexical-semantic candidate (apple) with equal levels of precision. This is different from the case in which predictions have to be derived from a sentence context, where the level of precision about the upcoming stimulus relies on the inferences made by the individual, based on contextual information and their prior knowledge. In the original study (Arcara et al., 2019), precision levels and the associated pre-stimulus alpha power was not manipulated; however, our question in the present study was whether spontaneous, physiological fluctuations of alpha power - regardless of experimental manipulation - have an influence on subsequent ERPs: we hypothesized that spontaneous alpha power fluctuations taking place on a trial-by-trial basis represented random variations in the ability to precisely predict the upcoming target word (Cao et al., 2017). In particular, lower levels of pre-stimulus alpha power should correspond to the ability to hold more precise expectations about the target, and when these expectations are disconfirmed in the case of incongruent targets, a N400 ensues, signaling linguistic/semantic prediction errors.

In spite of these differences, our results (albeit more nuanced) are in line with those from sentence comprehension paradigms. We have demonstrated that even if the N400 effect is observable regardless of pre-stimulus alpha power level, it is more negative for lower levels of pre-stimulus alpha power (Figure 4 and upper panels of Figure 5A-D). This suggests that prediction errors are only partly dependent on the pre-stimulus precision level, and that even in absence of explicit manipulations of pre-stimulus power or precision levels, some spontaneous variability exists, and may reflect variations in the ability to precisely predict the upcoming target word (Cao et al., 2017), so that in the presence of lower pre-stimulus alpha power, predictions about the upcoming targets are more precise, and generate a larger N400 (representing prediction error) following incongruent (i.e., unpredicted) targets. We hypothesize that this mechanism is similar to the one that takes place in those sentence comprehension paradigms that manipulate contextual constraint and find that more negative N400s follow context associated with lower pre-stimulus alpha and incongruent targets (Rommers et al., 2017; Terporten et al., 2019; Wang et al., 2018).

After a prediction error is generated, the cognitive system attempts to resolve it through a process of reanalysis, which is reflected in late posterior positivities/P600s (Kuperberg et al., 2020; Wang et al., 2021). The reanalysis process might be associated with changes in precision weights with respect to the precision configuration of the pre-stimulus interval: input reanalysis and prediction error resolution may be therefore dependent on and modulated by the precision level that was driving expectations in the pre-stimulus interval. However, these are only speculative explanations, as we did not manipulate the priors' precision levels.

As stated in Section 1.2, the attentional-based and precision-based interpretations of these electrophysiological dynamics are not in conflict with one another and are both compatible with a predictive account of the N400 and late posterior positivity/P600 effects. The predictive account also offers an overarching explanation of several other domains of cognition, that transcends the limitations of discrete cognitive modules and tends toward new ontologies (Pessoa et al., 2022; Poldrack and Yarkoni, 2016). In this vein, recent works are going in the direction of interpreting traditional cognitive functions under predictive views

(Masina et al., 2022) with promising implications for basic research and clinical applications. Notably, sometimes predictive processing theories, albeit increasingly popular, are criticized as lacking empirical support (Litwin and Milkowski, 2020; Wilterson et al., 2021). One of the reasons for this criticism is that studies explaining their results in terms of anticipation and prediction seldom take into account neurophysiological dynamics taking place before the stimuli, therefore assuming that predictive processes are taking place but not directly investigating them, while prediction is an ongoing process (Leon-Cabrera et al., 2017); investigations on this topic thus cannot disregard pre-stimulus dynamics. In the present study, we have established some evidence that pre-stimulus alpha power differently biases post-stimulus language-related ERPs, a phenomenon that can be parsimoniously explained by considering predictive processing in language. Deepening the understanding of the relationship between pre and post-stimulus neurophysiological dynamics, i.e., prediction formation and (dis)confirmation and their relative precision could be useful to settle the matter in favor of or opposing predictive processing theories. In this direction, future studies could manipulate the level of pre-stimulus attention and priors' precision, to establish the relative role of these variables; also, investigating the activity of different brain regions and networks in the pre vs post-stimulus interval using neuroimaging methods with a high spatial and temporal resolution could prove useful to disentangle between different neurocognitive mechanisms involved in linguistic anticipation and modulation of pre-stimulus alpha power and post-stimulus ERPs.

Limitations and future directions. There are some limitations in the present study. First of all, we acknowledge that the predictive and the non-predictive accounts are difficult to disentangle. Moreover, the study design was purposely simple so that it was possible to uncover pre- and post-stimulus dynamics specifically relative to very clear semantic violations, not embedded in a rich linguistic context. Future investigations should address this issue, performing single-trial analyses on more complex linguistic violation paradigms (e.g., sentence reading), to provide a more complete picture of the influence of the pre-stimulus state on later cognitive processes in linguistic comprehension. In the current study, the behavioral task was devised to ensure that the participants paid attention to the stimuli and was very easy and without time pressure. Future studies could employ different tasks that could link prestimulus alpha not only to EEG responses, but also to behavioral performance.

Finally, it is important to notice a potential limitation, which is intrinsic to most ERP studies on language: it is very difficult to demonstrate that the trial-by-trial alpha variations are actual spontaneous fluctuations and are not instead related to some uncontrolled properties of the stimuli. In short, since all stimuli were not equal, it could be the case that some specific cue-target pairs were associated to lower prestimulus alpha and in turn to enhanced N400 and P600. To explore this issue, we fitted additional GAMM models and ran control analyses based on simulations that can be found in the Supplementary Materials (p. 241). These additional analyses demonstrated that the effects we report seems not due to uncontrolled properties of the stimuli.

Conclusion. In the present study, we showed that ERPs associated with semantic violations, such as the N400 and late posterior positivity/P600, were differently modulated by pre-stimulus alpha power: the N400 amplitude was less sensitive to pre-stimulus power level (being present regardless of the level of pre-stimulus power), while larger late posterior positivities/P600s were only associated with low pre-stimulus alpha power level. More specifically, lower levels of pre-stimulus alpha power were associated with more negative N400 and more positive P600 ERPs. This finding can be explained through the different, albeit connected, functional meanings of the ERPs both under a predictive and a non-predictive theoretical framework. Capitalizing on the methodology proposed in the current study, further investigation of pre-stimulus dynamics could better characterize how pre-stimulus dynamics could affect brain responses.

Data Availability

Data analysis code is available at the OSF link https://osf.io/6jrk2/?view_only=00d6b579199a47758d8b8a2b854dabcf

Funding

GA and RP were funded by the Italian Ministry of Health. SL was funded by the Italian Ministry of Health and the University of Padua (Padua Neuroscience Center). SG was supported by a research fellowship on the grant PRIN20177894ZH funded by the Ministero della Ricerca awarded to FP.

References

Alamia, A., & VanRullen, R. (2019). Alpha oscillations and travelling waves: Signatures of predictive coding? *PLoS Biology*, *17*(10). https://doi.org/10.1101/464933

Alday, P. M., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2017). Electrophysiology reveals the neural dynamics of naturalistic auditory language processing: Event-related potentials reflect continuous model updates. *ENeuro*, *4*(6). https://doi.org/10.1523/ENEURO.0311-16.2017

Arcara, G., Franzon, F., Gastaldon, S., Brotto, S., Semenza, C., Peressotti, F., & Zanini, C. (2019). One can be some but some cannot be one : ERP correlates of numerosity incongruence are different for singular and plural. *CORTEX*, *116*, 104–121. https://doi.org/10.1016/j.cortex.2018.10.022

Arcara, G. & Petrova, A. erpR: Event-related potentials (ERP) analysis, graphics and utility functions (R version 0.2.0). https://rdrr.io/cran/erpR/ (2017).

Arcara, G., Pezzetta, R., Benavides-Varela, S., Rizzi, G., Formica, S., Turco, C., Piccione, F., & Semenza, C. (2021). Magnetoencephalography reveals differences in brain activations for fast and slow responses to simple multiplications. *Scientific Reports*, *11*(1), 1–13. https://doi.org/10.1038/s41598-021-97927-8

Baayen, R. H., Kuperman, V., Bertram, R., & Baayen, R. (2010). Frequency effects in compound processing. Compounding, Amsterdam/Philadelphia: Benjamins, 257-270.

Baayen, R. H., van Rij, J., de Cat, C., & Wood, S. (2018). *Autocorrelated Errors in Experimental Data in the Language Sciences: Some Solutions Offered by Generalized Additive Mixed Models*. 49–69. https://doi.org/10.1007/978-3-319-69830-4_4

Baayen, R. H., Vasishth, S., Bates, D., & Kliegl, R. (2017). Out of the cave of shadows. *Journal of Memory and Language*, *94*, 206–234. https://doi.org/10.1016/j.jml.2016.11.006

Baayen, R. H., Vasishth, S., Kliegl, R., & Bates, D. (2017). The cave of shadows: Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language*, *94*, 206–234. https://doi.org/10.1016/j.jml.2016.11.006

Bakeman, R. (2005). Recommended effect size statistic. *Behavior Research Methods*, *37*(3), 379–384. https://doi.org/10.3758/BF03192707

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. https://doi.org/10.1523/JNEUROSCI.3474-13.2014

Bazanova, O. M., & Vernon, D. (2014). Interpreting EEG alpha activity. *Neuroscience* *and Biobehavioral Reviews*, *44*, 94–110. https://doi.org/10.1016/j.neubiorev.2013.05.007

Bornkessel-Schlesewsky, I., Sharrad, I., Howlett, C. A., Alday, P. M., Corcoran, A. W., Bellan, V., Wilkinson, E., Kliegl, R., Lewis, R. L., Small, S. L., & Schlesewsky, M. (2022). Rapid adaptation of predictive models during language comprehension: Aperiodic EEG slope, individual alpha frequency and idea density modulate individual differences in real-time model updating. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.817516

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in Psychology*, *10*(FEB), 1–17. https://doi.org/10.3389/fpsyg.2019.00298

Brothers, T., Wlotko, E. W., Warnke, L., & Kuperberg, G. R. (2020). Going the Extra Mile: Effects of Discourse Context on Two Late Positivities During Language Comprehension. *Neurobiology of Language*, *1*(1), 135–160. https://doi.org/10.1162/nol_a_00006

Cao, L., Thut, G., & Gross, J. (2017). The role of brain oscillations in predicting self-generated sounds. *NeuroImage*, *147*(November 2016), 895–903. https://doi.org/10.1016/j.neuroimage.2016.11.001

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Cohen, J. (1983). The Cost of Dichotomization. *Applied Psychological Measurement*, *7*(3), 249–253. http://conservancy.umn.edu/bitstream/handle/11299/107497/v07n3p249.pdf?sequence=1

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

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

De Cat, C., Klepousniotou, E., & Baayen, R. H. (2015). Representational deficit or processing effect? An electrophysiological study of noun-noun compound processing by very advanced L2 speakers of English. *Frontiers in Psychology*, *6*(FEB), 1–17. https://doi.org/10.3389/fpsyg.2015.00077

Ergenoglu, T., Demiralp, T., Bayraktaroglu, Z., Ergen, M., Beydagi, H., & Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Cognitive Brain Research*, *20*(3), 376–383. https://doi.org/10.1016/j.cogbrainres.2004.03.009

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in human neuroscience*, *4*, 215.

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*(JUL), 1–13. https://doi.org/10.3389/fpsyg.2011.00154

Friston, K. (2012). Prediction, perception and agency. *International Journal of Psychophysiology*, *83*(2), 248–252. https://doi.org/10.1016/j.ijpsycho.2011.11.014

Friston, K. (2018). Does predictive coding have a future? *Nature Neuroscience*, *21*(8), 1019–1021. https://doi.org/10.1038/s41593-018-0200-7

Friston, K. J. (2019). Waves of prediction. *PLoS Biology*, *17*(10), 1–7. https://doi.org/10.1371/journal.pbio.3000426

Garrido, M. I., & Deouell, L. Y. (2021). Unilateral neglect within the predictive processing framework. *Brain Communications*, *3*(3), 1–4. https://doi.org/10.1093/braincomms/fcab193

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

Gastaldon, S., Arcara, G., Navarrete, E., & Peressotti, F. (2020). Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. *Cortex*, *133*, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026

Gastaldon, S., Busan, P., Arcara, G., Peressotti, F. (2023). Inefficient speech-motor control affects predictive speech comprehension: atypical electrophysiological correlates in stuttering. *Cerebral Cortex*. DOI: https://doi.org/10.1093/cercor/bhad004

Grent-'T-Jong, T., Boehler, C. N., Kenemans, J. L., & Woldorff, M. G. (2011). Differential functional roles of slow-wave and oscillatory-alpha activity in visual sensory cortex during anticipatory visual-spatial attention. *Cerebral Cortex*, *21*(10), 2204–2216. https://doi.org/10.1093/cercor/bhq279

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K. H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *NeuroImage*, *37*(4), 1465–1473. https://doi.org/10.1016/j.neuroimage.2007.07.011

Hodapp, A., & Rabovsky, M. (2021). The N400 ERP component reflects an error□ based implicit learning signal during language comprehension. *European Journal of Neuroscience*, *54*(9), 7125-7140.

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, *31*(1), 19–31. https://doi.org/10.1080/23273798.2015.1072223

Hutchinson, B. T., Pammer, K., & Jack, B. (2021). Pre-stimulus alpha predicts inattentional blindness. *Consciousness and Cognition*, *87*(September 2020), 103034. https://doi.org/10.1016/j.concog.2020.103034

Iemi, L., Busch, N. A., Laudini, A., Haegens, S., Samaha, J., Villringer, A., & Nikulin, V. V. (2019). Multiple mechanisms link prestimulus neural oscillations to sensory responses. *ELife*, *8*(e43620), 1–34. https://doi.org/10.1101/461558

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, *4*(November), 1–8. https://doi.org/10.3389/fnhum.2010.00186

Jiang, J., Summerfield, C., & Egner, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *Journal of Neuroscience*, *33*(47), 18438–18447. https://doi.org/10.1523/JNEUROSCI.3308-13.2013

Katyal, S., He, S., He, B., & Engel, S. A. (2019). Frequency of alpha oscillation predicts individual differences in perceptual stability during binocular rivalry. *Human Brain Mapping*, *40*(8), 2422–2433. https://doi.org/10.1002/hbm.24533

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A Tale of Two Positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, *32*(1), 12–35. https://doi.org/10.1162/jocn

Kutas, M., & Hillyard, S. A. (1984). Brain Potentials During Reading Reflect Word Expectancy and Semantic Association. *Nature*, *307*(5947), 161–163.

Leckey, M., & Federmeier, K. D. (2020). The P3b and P600(s): Positive contributions to language comprehension. *Psychophysiology*, *57*(7), 1–15. https://doi.org/10.1111/psyp.13351

Litwin, P., & Miłkowski, M. (2020). Unification by Fiat: Arrested Development of Predictive Processing. *Cognitive Science*, *44*(7). https://doi.org/10.1111/cogs.12867

León-Cabrera, P., Piai, V., Morís, J., & Rodríguez-Fornells, A. (2022). Alpha power decreases associated with prediction in written and spoken sentence comprehension. *Neuropsychologia*, 108286. https://doi.org/10.1016/j.neuropsychologia.2022.108286

León-Cabrera, P., Rodríguez-Fornells, A., & Morís, J. (2017). Electrophysiological correlates of semantic anticipation during speech comprehension. *Neuropsychologia*, *99*, 326–334. https://doi.org/10.1016/j.neuropsychologia.2017.02.026

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, *68*, 155–168. https://doi.org/10.1016/j.cortex.2015.02.014

Limbach, K., & Corballis, P. M. (2016). Prestimulus alpha power influences response criterion in a detection task. *Psychophysiology*, *53*(8), 1154–1164. https://doi.org/10.1111/psyp.12666

Luck, S. J. (2005). Ten Simple Results of Designing ERP Experiments. In Event-Related Potentials A Methods Handbook, 4.

Luck, S. J. (2014). Filtering as a Time-Domain Procedure. In Gazzaniga, M. S., *An Introduction to the Event-Related Potential Technique* (2nd edition, pp. 193-200). MIT press.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40.

https://doi.org/10.1037/1082-989X.7.1.19

Mancini, S., Massol, S., Duñabeitia, J. A., Carreiras, M., & Molinaro, N. (2019). Agreement and illusion of disagreement: An ERP study on Basque. *Cortex*, *116*, 154–167. https://doi.org/10.1016/j.cortex.2018.08.036

Mantegna, F., Hintz, F., Ostarek, M., Alday, P. M., & Huettig, F. (2019). Distinguishing integration and prediction accounts of ERP N400 modulations in language processing through experimental design. *Neuropsychologia*, *134*(September), 107199. https://doi.org/10.1016/j.neuropsychologia.2019.107199

Masina, F., Pezzetta, R., Lago, S., Mantini, D., Scarpazza, C., & Arcara, G. (2022). Disconnection from prediction: A systematic review on the role of right temporoparietal junction in aberrant predictive processing. *Neuroscience & Biobehavioral Reviews*, *138*(January), 104713. https://doi.org/10.1016/j.neubiorev.2022.104713

Meulman, N., Wieling, M., Sprenger, S. A., Stowe, L. A., & Schmid, M. S. (2015). Age Effects in L2 Grammar processing as revealed by ERPs and How (Not) to Study Them. *PLoS ONE*, *10*(12), 1–27. https://doi.org/10.1371/journal.pone.0143328

Min, B. K., & Herrmann, C. S. (2007). Prestimulus EEG alpha activity reflects prestimulus top-down processing. *Neuroscience Letters*, *422*(2), 131–135. https://doi.org/10.1016/j.neulet.2007.06.013

Nalaye, H., Cross, Z. R., Schlesewsky, M., & Bornkessel-schlesewsky, I. (2022). Electrophysiological indices of individual differences in adult language learning. *BioRxiv*, 1–43.

Nicenboim, B., Vasishth, S., & Rösler, F. (2020). Are words pre-activated probabilistically during sentence comprehension? Evidence from new data and a bayesian random-effects meta-analysis using publicly available data. *Neuropsychologia*, *142*(107427).

https://doi.org/10.1016/j.neuropsychologia.2020.107427

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., Ferguson, H. J., Fu, X., Heyselaar, E., Huettig, F., Husband, E. M., Ito, A., Kazanina, N., Kogan, V., Kohút, Z., Kulakova, E., Mézière, D., Politzer-Ahles, S., Rousselet, G., … Von Grebmer Zu Wolfsthurn, S. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a largescale study using brain potentials. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(20180522). https://doi.org/10.1098/rstb.2018.0522

Pessoa, L., Medina, L., & Desfilis, E. (2022). Refocusing neuroscience: moving away from mental categories and towards complex behaviours. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *377*(1844), 20200534. https://doi.org/10.1098/rstb.2020.0534

Poldrack, R. A., & Yarkoni, T. (2016). From brain maps to cognitive ontologies: informatics and the search for mental structure. *Annual Review of Psychology*, *67*(January 4), 587–612. https://doi.org/10.1146/annurev-psych-122414-033729.From

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Porretta, V., Tremblay, A., & Bolger, P. (2017). Got experience? PMN amplitudes to foreign-accented speech modulated by listener experience. *Journal of Neurolinguistics*, *44*, 54–67. https://doi.org/10.1016/j.jneuroling.2017.03.002

Price, G. W. (1997). The Effect of Pre-Stimulus Alpha Activity on the Auditory P300 Paradigm: A Prospective Study. *Brain Topography*, *9*(3), 169–176. https://doi.org/10.1111/j.1469-8986.1988.tb00979.x

Quante, L., Bölte, J., & Zwitserlood, P. (2018). Dissociating predictability, plausibility and possibility of sentence continuations in reading: Evidence from late-positivity ERPs.

PeerJ, *2018*(10), 1–24. https://doi.org/10.7717/peerj.5717

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rommers, J., Dickson, D. S., Norton, J. J. S., Wlotko, E. W., & Federmeier, K. D. (2017). Alpha and theta band dynamics related to sentential constraint and word expectancy. *Language, Cognition and Neuroscience*, *32*(5), 576–589. https://doi.org/10.1080/23273798.2016.1183799

Samaha, J., Boutonnet, B., Postle, B. R., & Lupyan, G. (2018). Effects of meaningfulness on perception: Alpha-band oscillations carry perceptual expectations and influence early visual responses. *Scientific Reports*, *8*(1), 1–14. https://doi.org/10.1038/s41598-018-25093-5

Sassenhagen, J., & Bornkessel-Schlesewsky, I. (2015). The P600 as a correlate of ventral attention network reorientation. *Cortex*, *66*, A3–A20. https://doi.org/10.1016/j.cortex.2014.12.019

Sassenhagen, J., & Fiebach, C. J. (2019). Finding the P3 in the P600: Decoding shared neural mechanisms of responses to syntactic violations and oddball targets. *NeuroImage*, *200*(May), 425–436. https://doi.org/10.1016/j.neuroimage.2019.06.048

Sedley, W., Gander, P. E., Kumar, S., Kovach, C. K., Oya, H., Kawasaki, H., Howard, M. A., & Griffiths, T. D. (2016). Neural signatures of perceptual inference. *ELife*, *5*(MARCH2016), 1–13. https://doi.org/10.7554/eLife.11476

Sherman, M. T., Kanay, R., Seth, A. K., & VanRullen, R. (2016). Rhythmic Influence of Top–Down Perceptual Priors in the Phase of Prestimulus Occipital Alpha Oscillations. *Journal of Cognitive Neuroscience*, *28*(9), 1318–1330. https://doi.org/10.1162/jocn_a_00973

Sóskuthy, M. (2021). Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics*, *84*. https://doi.org/10.1016/j.wocn.2020.101017

Sulpizio, S., Arcara, G., Lago, S., Marelli, M., Amenta, S. (*in press*). Very early and late form-to-meaning computations during visual word recognition as revealed by electrophysiology. *Cortex*.

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational intelligence and neuroscience*, *2011*, 879716. https://doi.org/10.1155/2011/879716

Terporten, R., Schoffelen, J. M., Dai, B., Hagoort, P., & Kösem, A. (2019). The Relation between Alpha/Beta Oscillations and the Encoding of Sentence induced Contextual Information. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-56600-x

Tremblay, A., Asp, E., Johnson, A., Migdal, M. Z., Bardouille, T., & Newman, A. J. (2016). What the Networks Tell us about Serial and Parallel Processing: An MEG Study of Language Networks and N-gram Frequency Effects in Overt Picture Description. *The Mental Lexicon*, *11*(1), 115–160.

Tremblay, A., & Newman, A. J. (2015). Modeling nonlinear relationships in ERP data using mixed-effects regression with R examples. *Psychophysiology*, *52*, 124–139. https://doi.org/10.1111/psyp.12299

Tsiwah, F., Bastiaanse, R., van Rij, J., & Popov, S. (2021). Online Processing of Temporal Agreement in a Grammatical Tone Language: An ERP Study. *Frontiers in Psychology*, *12*(May), 1–15. https://doi.org/10.3389/fpsyg.2021.638716

Van De Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwill, D. J. (2010). Monitoring in language perception: Mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience*, *22*(1), 67–82.

https://doi.org/10.1162/jocn.2008.21170

Van Den Berg, B., Appelbaum, L. G., Clark, K., Lorist, M. M., & Woldorff, M. G. (2016). Visual search performance is predicted by both prestimulus and poststimulus electrical brain activity. *Scientific Reports*, *6*(May), 1–13. https://doi.org/10.1038/srep37718

Van Diepen, R. M., Foxe, J. J., & Mazaheri, A. (2019). The functional role of alpha-band activity in attentional processing: the current zeitgeist and future outlook. *Current Opinion in Psychology*, *29*, 229–238. https://doi.org/10.1016/j.copsyc.2019.03.015

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

van Rij, J., Hendriks, P., van Rijn, H., Baayen, R. H., & Wood, S. N. (2019). Analyzing the Time Course of Pupillometric Data. *Trends in Hearing*, *23*, 1–22. https://doi.org/10.1177/2331216519832483

van Rij J, Wieling M, Baayen R, van Rijn H (2020). “itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs.” R package version 2.4.

Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. Annals of the new York Academy of Sciences, 1464(1), 242-268.

Wang, P., He, Y., Maess, B., Yue, J., Chen, L., Brauer, J., Friederici, A. D., & Knösche, T. R. (2022). Alpha power during task performance predicts individual language comprehension. NeuroImage, 260(119449). https://doi.org/10.1016/j.neuroimage.2022.119449

Wang, L., Schoot, L., Alexander, E., Warnke, L., Kim, M., Khan, S., Hamalainen, M. S., & Kuperberg, G. R. (2021). Dynamic predictive coding across the left fronto-temporal language hierarchy: Evidence from MEG, EEG and fMRI. *BioRxiv*, 1–55.

Wang, L., Hagoort, P., & Jensen, O. (2018). Language Prediction Is Reflected by Coupling between Frontal Gamma and Posterior Alpha Oscillations. *Journal of Cognitive Neuroscience*, *30*(3), 432–447.

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, *70*, 86–116. https://doi.org/10.1016/j.wocn.2018.03.002

Wilterson, A. I., Nastase, S. A., Bio, B. J., Guterstam, A., & Graziano, M. S. A. (2021). Attention, awareness, and the right temporoparietal junction. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(25). https://doi.org/10.1073/pnas.2026099118

Wood, S. N. (2006). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

Interim summary 2

So far, Study 1 showed that rTPJ is a possible domain-independent predictive hub, while Study 2 highlighted the importance of considering the pre-stimulus interval when studying predictive dynamics, and showed that the level of pre-stimulus alpha power can modulate ERP amplitudes. To further elucidate the issue of TPJs in predictive processing Study 3 will combine the investigation of pre-stimulus dynamics with that of the role of bilateral TPJs during complex linguistic computations taking place during a metaphor comprehension task, a particular case of linguistic predictive processing (Vespignani et al., 2010). More specifically, this study attempts to answer the following questions: are *both* TPJs involved in linguistic prediction? Are they involved in both prediction generation and testing, as proposed by Siman-Tov et al. (2019), or only in prediction testing, as proposed by Doricchi et al. (2022)? If the TPJs are involved in prediction generation, can pre-stimulus alpha (associated with predictions' precision) modulate the subsequent brain responses to target stimuli? And finally, is the eventual modulation local, i.e., limited to the TPJs under investigation, or can pre-stimulus TPJ activity influence post-stimulus activations in other task-related areas? The results of this study can help shed some light on the possibly different roles of left and right TPJ during complex linguistic computations, and at what stage of predictive processing they come into play.

Study 3: Roles of left and right TPJ during linguistic predictive processing

1. Introduction

1.1. The many facets of the TPJs

The bilateral TPJs have been found to play a role in a variety of cognitive processes, for example, attention, language, body ownership and sense of agency, and episodic memory retrieval. Bilateral TPJs are also part of the default mode network (DMN), which is active during social cognition tasks (see General Introduction). However, the left and right TPJs also show some specificity in their functional roles. For example, attentional tasks seem to activate more specifically the right TPJ (rTPJ), especially in response to unexpected but task-relevant stimuli (e.g., Doricchi et al., 2010; Indovina & Macaluso, 2007). This leads to the hypothesis that rTPJ could play a key role in the reorienting of attention (Corbetta et al., 2008; Corbetta & Shulman, 2002). This hypothesis, known as the circuit-breaking theory, states that the dorsal attention network (DAN) maintains the visuospatial information relevant to the current task-defined goals, while the ventral attention network (VAN), including rTPJ, allows the switch of attention to relevant but currently unattended stimuli; rTPJ would therefore be responsible for interrupting the activity of the DAN, resulting in the reorientation of attention to a new salient stimulus (Corbetta et al., 2008; Corbetta & Shulman, 2002).

On the other side the left TPJ (lTPJ) appears to be more involved in language processing: it is activated in response to lexical violations, non-words, and words that are semantically unrelated to the previous context (Binder et al., 2005, p. 200; Fiebach et al., 2002; Prince et al., 2007), but it is also in charge of integrating the individual's general world knowledge with the local discourse information (Menenti et al., 2008; Metusalem et al., 2012) and in pre-activating linguistic information (Gastaldon et al., 2020). In addition, some studies link lTPJ to figurative language comprehension (Bambini et al., 2011), but there is also evidence that this type of task activates the TPJs bilaterally (Bohrn et al., 2012; Spotorno et al., 2012).

The involvement of the bilateral TPJs in such heterogeneous contexts has led to some more domain-general conceptualizations of the role of these areas (e.g., nexus model, Carter & Huettel, 2013; contextual updating hypothesis, Geng & Vossel, 2013; see General Introduction) stating that rTPJ in particular would act as a hub for integrating information from multiple domains and thus updating the internal models of the world.

Internal models are particularly important in predictive processing theories, that have been extensively presented in the General Introduction and in Study 2. Predictive computations are believed to be carried out in a distributed way across the brain: a recent meta-analysis proposed that a set of brain areas, comprising also rTPJ, form a diffuse network involved in higher-level prediction generation and testing, that supports both perception and action processes in a domain-general way (Siman-Tov et al., 2019).

Besides the putative prediction network, rTPJ is part of both the VAN and the DMN (Corbetta et al., 2008; Hughes et al., 2019), a concurrency that further speaks in favor of a domain-general conceptualization of this area as a hub where multiple cognitive processes and information types converge and are integrated, as reported in Study 1. In essence, Study 1 concludes that the role of rTPJ is driven by the specific context and network it operates in, while its core function of integration and contextual updating remains unchanged. This inherent flexibility of rTPJ is likely why it is found to be involved across various cognitive domains, as it serves a fundamental function in a range of cognitive processes (Masina et al., 2022).

These findings support an interpretation of the role of rTPJ in terms of predictive processing. In contrast to traditional views that emphasize the involvement of rTPJ in distinct domains, this evidence lays the case for a more parsimonious interpretation of the role of the rTPJ (and, probably, also lTPJ) as a crucial hub involved in domain-general predictions, together with other areas.

1.2. Metaphors: a case of linguistic predictive processing

Predictive processing in the language domain is being extensively and increasingly studied in recent years (e.g., Brown & Kuperberg, 2015; Kuperberg et al., 2020); notably, the figurative use of language can be intended as a special case of linguistic prediction (Vespignani et al., 2010). Figurative language includes various forms such as metaphors, idioms, and irony, all situations where the intended meaning goes beyond the literal content of the single sentence components (Weiland et al., 2014). For instance, metaphors offer concise and evocative descriptions by creatively connecting two different concepts (e.g., “that lawyer is a shark”). In sum, differently from literal expressions, figurative ones include both literal and non-literal meanings (Diaz & Eppes, 2018), and the principles of predictive processing can help us understand how we make sense of figurative language.

Goodman and Frank (2016) argued that speakers produce sentences that are both helpful and parsimonious, relative to some particular topic or goal. Comprehenders then understand these sentences by inferring what the speaker must have meant, given what she said. This means that comprehenders infer the underlying cause (what the speaker intends to communicate) based on the sensory input (what the speaker said). Expanding on this observation, we can hypothesize that in the case of metaphors, the sensory input can have more than one meaning: it can be both literally and metaphorically true (Bohrn et al., 2012), but to correctly interpret the intended meaning of the sentence, i.e., inferring whether the metaphorical or the literal meaning is the true one, we have to rely on some expectations coming from our prior knowledge. Consider the sentence “that lawyer is a shark”. While reading it, we generate a hypothesis about the general meaning of the sentence based on our prior knowledge about lawyers. We are, in other words, *uncertain* about the plausible implied meaning of the sentence (Brown & Kuperberg, 2015), but we have some priors, coming from our general knowledge, that can help us in the interpretation. Thanks to these priors we can understand that the literal meaning makes no sense, because it is implausible that the correct meaning of the sentence is that the lawyer is a predatory fish that swims underwater. Therefore, we have to infer that other features of the concept of “shark”, for example, its ferocity or aggressiveness, constitute the correct implied meaning of the sentence. Another feature guiding the interpretation of the sentence is the context, which can shape our predictions. For example, if the sentence is uttered by a young boy that is watching a cartoon, it could actually describe a shark that works as a lawyer. During predictive linguistic computations, the degree of uncertainty (i.e., *precision*) associated with our prediction about the meaning can vary, but usually our priors, built from general knowledge and and context, guide us through the inference with a fair degree of precision and are implied in updating our internal models of the sentence. The concept of precision mathematically represents the reliability of a prediction, and is defined as the prediction’s inverse variance (Friston, 2018), even though it can also be associated with perceived stimuli or prediction errors (Walsh et al., 2020). In essence, the higher the precision of a stimulus or expectation, the more the individual relies on it; therefore, highly precise expectations reduce the uncertainty associated with the underlying cause of upcoming stimuli (in the case of figurative language comprehension, about the more plausible meaning of a metaphor).

During linguistic prediction, our internal models test the possible sentence meaning against our prior knowledge of the world and the actual sensory input until we come up with a plausible interpretation. This requires integrating global world knowledge and local sentence information, and unsurprisingly, beyond more classical linguistic areas, lTPJ seems to be particularly implicated in such a process (Doricchi et al., 2022; Gastaldon et al., 2020; Menenti et al., 2008; Metusalem et al., 2012), but there is also evidence showing involvement of rTPJ during metaphor (Bambini et al., 2011) and, in general, figurative language comprehension (Spotorno et al., 2012).

1.3. Are TPJs involved in both prediction generation and testing?

A recent systematic literature review proposed that the lTPJ encodes both matches and mismatches between predicted and observed sensory, motor, or cognitive events, whereas the rTPJ specifically encodes mismatches; and that match/mismatch computations follow the prediction error minimization principle (Doricchi et al., 2022). Following this line of reasoning, we can argue that both TPJs are active after the encounter with the actual environmental stimuli, and they are involved in checking whether the stimuli are or not in line with the predictions, i.e., in the prediction testing phase. Conversely, Siman-Tov et al. (2019) presume that their putative prediction network, encompassing TPJ, is involved in both generating and testing predictions; this would mean that TPJ is active both before and after encountering the crucial environmental stimuli. However, the studies included in their meta-analysis and those reviewed so far have only detected TPJ activation *after* predictions were disconfirmed, only assuming that the detected mismatch was the consequence of a violated prediction, but not directly investigating whether a prediction is actually taking place. As a consequence, we can argue that the TPJs are involved in prediction violation, rather than in prediction *per se*. Hence, some questions arise: is also lTPJ involved in prediction? Are the TPJs only involved in prediction testing or also in prediction generation? Answering these questions implies investigating the TPJs' activity during cognitive tasks, not only after target stimuli but also before them. We, therefore, have to direct our interest toward the TPJs' state during the pre-stimulus interval, which could be representative of prediction generation, and, as such, might modulate the neural responses to subsequent stimuli.

1.3.1. Alpha, precision, and attention

To the best of our knowledge, no previous studies have directly investigated the influence of spontaneous (i.e., in the absence of manipulation) pre-stimulus TPJ activity on the post-stimulus processing. Still, some preliminary evidence from EEG studies shows that different levels of pre-stimulus alpha power (8-13 Hz) might modulate post-stimulus ERPs: in simple perceptual paradigms, lower levels of pre-stimulus alpha power have been linked to larger ERP amplitudes (Cao et al., 2017; Ergenoglu et al., 2004; Grent-'T-Jong et al., 2011; Samaha et al., 2018; Van Den Berg et al., 2016), while higher pre-stimulus alpha power has been associated with ERP amplitude suppression, longer latencies, and augmentation of late ERPs (occurring after 400 milliseconds; Iemi et al., 2017, 2019; Min & Herrmann, 2007). This evidence does not give us any insight into the role of the TPJs, but it suggests that pre-stimulus alpha power can influence stimulus processing and that its fluctuations can be interpreted as possible correlates of anticipation or predictive processing.

Alpha power has been traditionally considered indicative of anticipatory attentional suppression (Foxe & Snyder, 2011), controlling the redirection of information flow to brain areas relevant to the task at hand, while inhibiting irrelevant regions (Jensen & Mazaheri, 2010). Under this perspective, increases in alpha power before a target stimulus suggest a state of cortical inhibition, whereas decreases in alpha power indicate cortical activation, facilitating the subsequent detection and processing of task-relevant information (Klimesch et al., 2007). In sum, during a state of anticipatory attention or cortical activation, stimulus processing should be facilitated; in this context, an unexpected stimulus would elicit an enhanced ERP.

On the other hand, a predictive conceptualization of alpha oscillations proposes that they are linked to the concept of precision (see Section 1.3; Bauer et al., 2014; Sedley et al., 2016; Sherman et al., 2016). Notably, in the context of predictive processing theories, attention is regarded as an emergent property of the precision optimization mechanism, where directing attention to a stimulus involves representing and enhancing the precision of sensory information (including prediction error) throughout the inferential process (Feldman & Friston, 2010). As mentioned above, highly precise expectations reduce the uncertainty associated with the underlying cause of upcoming stimuli (in the previously described case of metaphor comprehension, about the more plausible meaning); as a consequence, the relative internal representations will be enhanced and so will the neural responses associated with them. In this way, cognitive resources will be biased towards more precise predictions (Hohwy, 2012), and the magnitude of expectations' effects on neural activity depends on the precision of predictions and prediction error signals (Walsh et al., 2020). Within this theoretical framework, alpha oscillations are considered correlates of the prior's precision and capable of modulating post-stimulus neural responses, and there appears to be a correlation between highly precise (reliable) predictions and reduced pre-stimulus alpha power (Bauer et al., 2014; Cao et al., 2017). Conversely, elevated levels of pre-stimulus alpha power have been linked to dampened stimulus-evoked responses (Cao et al., 2017). It is however important to note that variations in pre-stimulus alpha power can also arise spontaneously, without any experimental manipulations; additionally, within a specific task, the features of the stimuli do not exhibit a linear relationship with the precision of their internal representations, as predictive models include random fluctuations of precision (state-dependent error variance) and assume that this precision is not constant at any level of hierarchical inference (Feldman & Friston, 2010). As a result, precision levels may fluctuate spontaneously during the task, similar to attention and alpha power.

This view on alpha oscillations aligns also with traditional "inhibitory" theories of attention, stating that anticipatory allocation of cognitive resources is reflected in lower levels of alpha power, facilitates stimulus processing and results in enhanced ERPs. Nevertheless, predictive processing differs from traditional accounts as it proposes a continuous interplay between attention and expectation that is aimed at optimizing precision and minimizing prediction error (Walsh et al., 2020). Interestingly enough, Siman-Tov et al. (2019) substantiate the idea that prediction and attention are interdependent processes: they showed that their putative prediction network and the VAN, subserving attentional reorienting, overlap in the rTPJ, and some of the conditions activating the VAN also involve violation of predictions (Corbetta et al., 2008; Vossel et al., 2014).

1.3.2. Pre-stimulus alpha in language

Prediction violation is an important feature when studying predictive processing, especially in the domain of language, in which expectations are mainly studied using EEG paradigms where the final target word is incongruent with the previous sentence contexts, and typically, incongruent (therefore unpredictable) words elicit characteristic N400s and P600s ERPs, associated with linguistic prediction error and its resolution, respectively (Kuperberg et al., 2020). Similarly to what happens for the TPJs, whose role in prediction has been inferred primarily based on their involvement after disconfirmation of expectations, the main drawback of linguistic violation paradigms is that they focus on the neural dynamics occurring after predictions have been either confirmed or disconfirmed; as a result, these experiments assume that prediction is occurring before the target, but they do not directly address it (León-Cabrera et al., 2017). To overcome this limitation, some studies on linguistic processing have considered both the pre- and post-stimulus interval, in different ways. What emerges is that more reliable, stronger linguistic expectations (induced, for example, by a more constraining context) were associated with a reduced pre-stimulus alpha power and more negative N400s (Gastaldon et al., 2023; Rommers et al., 2017; Terporten et al., 2019; Wang et al., 2018). In addition, in a semantic congruence task where image cues induced highly precise expectations about target words that could be congruent or incongruent, spontaneous fluctuations of pre-stimulus alpha power modulated the amplitude of the N400 and P600 components: the lower the pre-stimulus power, the greater the ERP amplitudes (Lago et al., 2023). Such an influence of fluctuations in pre-stimulus alpha power on the subsequent neural responses supports the hypothesis that alpha oscillations are associated with the prediction’s precision level, which modulates how the upcoming stimuli will be processed. Alpha oscillations can therefore be interpreted as variations in the individual’s ability to predict the upcoming target word with a high level of precision (Cao et al., 2017).

This mechanism could be coherent with the predictive processing of metaphors as outlined in section 1.3: linguistic predictions in favor of literal meanings, guided by prior world knowledge, might be associated with varying levels of precision and, therefore, with varying levels of pre-stimulus alpha power. Where and how in the brain this mechanism is carried out remains an open question. For example, linguistic prediction could only be carried out within the language network; or maybe these computations might also involve other domain-general predictive hubs, that are functionally connected with the language network and, probably, are also active in other cognitive operations (Ryskin & Nieuwland, in press).

As reviewed so far, there are reasons to hypothesize that, even if they are not part of the language network, the TPJs (or at least the left one) might come into play during linguistic prediction. If, as some studies report (Doricchi et al., 2022; Siman-Tov et al., 2019), the TPJs are involved in domain-general prediction generation and/or testing, and they are functionally connected with task-specific areas (Masina et al., 2022), we can hypothesize that pre-stimulus alpha power recorded in the TPJs (associated with the prior's precision), could modulate the neural responses to metaphor comprehension.

1.4. The present study

Metaphor comprehension can be regarded as a particular form of linguistic predictive processing, possibly related to priors' precision (see Sections 1.2. and 1.3.2.). In predictive frameworks, fluctuations in precision have been linked with pre-stimulus alpha oscillations, which can modulate the post-stimulus EEG responses (see Study 2, Lago et al., 2023); in addition, the prior's precision guides resources' allocation, probably with the involvement of rTPJ (the overlapping point between the VAN, a domain-specific network engaged in attentional processes, and a putative domain-general prediction network). Such evidence points to the importance of considering, when studying predictive dynamics, both precision variations and the potential functional connections between rTPJ and other areas. This last consideration is also valid for lTPJ, which has been somewhat overlooked in the investigations about predictive processing. In addition, the TPJs have been rarely studied in the context of linguistic prediction (e.g. Gastaldon, 2020), despite some evidence that rTPJ may be part of a predictive network while lTPJ might be involved in (linguistic) prediction by integrating local, task-related linguistic information with our existing world knowledge, with the likely aim of formulating the priors guiding figurative language comprehension (Menenti et al., 2008; Metusalem et al., 2012).

For these reasons, the present study aims at investigating the role of the bilateral TPJs in prediction generation and testing, during the comprehension of short literal and metaphoric sentences that participants were required to read; at the end of each sentence, they had to select, between two words, the one that was best related to the preceding phrase.

More specifically, we want to answer the questions outlined in Section 1: are *both* TPJs involved in linguistic prediction? Are they involved in both prediction generation and testing, as proposed by Siman-Tov et al. (2019), or only in prediction testing, as proposed by Doricchi et al. (2022)? If the TPJs are involved in prediction generation, can pre-stimulus alpha (associated with predictions' precision) modulate the subsequent brain responses to target stimuli? Finally, is the eventual modulation local, i.e., limited to the TPJs under investigation, or can pre-stimulus TPJ activity influence post-stimulus activations in other task-related areas?

These questions pose some methodological challenges that need to be addressed carefully. The majority of studies reviewed so far are fMRI studies (or meta-analyses and systematic literature reviews based on fMRI studies) with a very good spatial resolution but a limited temporal resolution, whereas the investigation of pre-stimulus TPJ dynamics requires adequate resolution in both temporal and spatial domains. Magnetoencephalography (MEG) meets these requirements, since it allows a reliable source reconstruction of the signal together with an excellent temporal resolution, enabling the investigation of the influence of the pre-stimulus oscillations on the post-stimulus signal both within the TPJs and in connection with other task-specific, language-related areas. As regards those areas that may be involved in the predictive processing of metaphoric sentences together with the TPJs, we chose to consider three core regions of interest (ROIs) within the language network, namely, Broca's area, and the superior and middle temporal gyri (STG and MTG; Friederici, 2017; Rapp et al., 2012).

In sum, the present study aimed to investigate whether the left and right TPJs could be functionally connected to other language-related regions, by testing whether pre-stimulus alpha recorded from the TPJs could predict post-stimulus activations both in the TPJs themselves and in the linguistic areas.

2. Method

The study was approved by the local ethics committee (Comitato Etico per la Sperimentazione Clinica della provincia di Venezia e IRCCS San Camillo) and conducted following the guidelines of the Declaration of Helsinki.

2.1. Participants

A sample of 28 healthy participants took part in the study. All participants had no neurological or psychiatric disease that could affect cognitive performance and had no history of developmental dyslexia or dyscalculia. The mean age of participants was 28.14 years (SD = 5.46, range = 21-45) and their mean education was 17.25 years (SD = 2.10, range = 13-21). Seventeen participants were female and 11 were male. All participants were right-handed.

2.2. Procedures

Before entering the magnetically shielded room, participants underwent initial preparation, which consisted of the placement of three head coils, to monitor head position during MEG recording, and eight external electrodes to record VEOG, HEOG, and ECG (all with bipolar montage). Before recordings coil positions and head shape were digitized using the Polhemus Fastrak system.

Continuous MEG signal was acquired using a whole-head 275-channel system (CTF-MEG). Data were sampled at 1200 Hz, with a hardware anti-aliasing low pass filter at 300 Hz. Before the experimental session, participants underwent a five-minute resting state session and then undertook the experimental task which consisted of four runs of about five minutes each. During the breaks between each run participants were allowed to rest and adjust their position. The total duration of the experiment was approximately 45 minutes. In all recordings, head movements never exceeded the threshold of 5 mm (on any axis).

All participants had also an MRI scan for source localization purposes (Philips Achieva, 1.5 T, T1-3d sequence). If the participant had already performed this examination in the institution and the data was available, we retrieved the scan. If not available, an MR scan was performed.

2.3. Materials

Participants were instructed that they would have to perform a language task that would consist of reading short word-by-word sentences and that at the end they would have to select, between two words, the one that was best related to the preceding sentence. There was no mention of the fact that the focus was on metaphors or pragmatics.

The task recording session was divided into four runs of about five minutes each, with small breaks between each run. The experiment was programmed with Psychopy (Peirce, 2007) (version 1.82), running on a PC.

Each trial was organized as follows: first, a fixation cross (+) was presented for 1500 ms, followed by the five words composing the sentence, lasting 300 ms, and presented each on a separate screen. Each word was preceded by a blank screen lasting 200 ms (e.g. “That/ lawyer/ is/ a/ shark”). A schematic representation of each trial can be found in Figure 1. After the whole sentence, a blank screen lasting from 1200 ms to 1600 ms was presented, followed by two words, one on the left and one on the right of the screen (e.g. “precise”, “aggressive”). Participants were instructed to respond as quickly and as accurately as possible by pressing one of two buttons with the index finger or the middle finger the word that was more closely related to the meaning of the preceding sentence. The hand to be used for the task was counterbalanced across participants and the buttons to press were always congruent to the spatial location of the words (left or right). Reaction Times and Accuracy were recorded for this button press. Stimuli were presented using white Courier New font on a black background screen, and all stimuli subtended about 1.5° of visual angle on the horizontal plane.

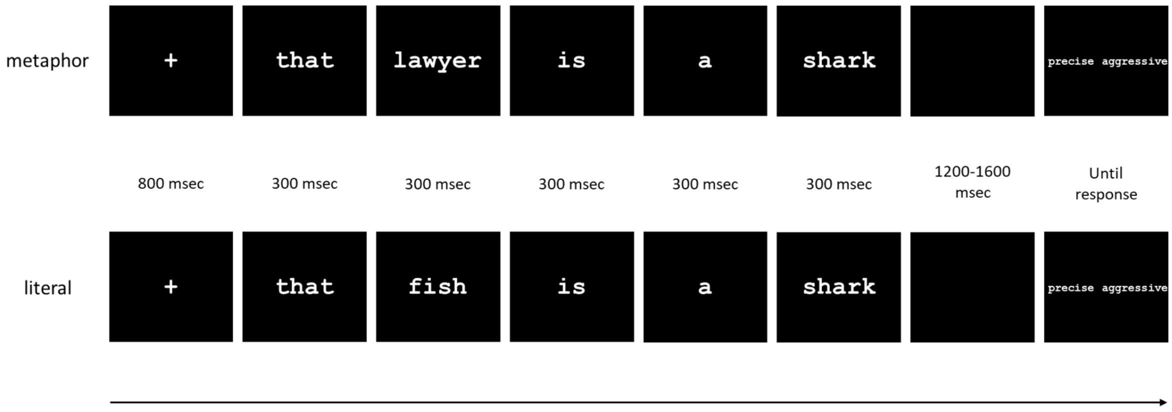


Figure 1. Examples of metaphorical and literal sentence stimuli. Sentences were preceded by a fixation cross (800 msec duration). Sentences were presented word-by-word and each word word remained on the screen for 300 msec. ISI lasted 200 msec, while the adjectives remained on the screen until participants gave their response.

Experimental items consisted of 164 sentences with the form “that X is a Y” (for similar stimuli see Bambini, Bertini, Schaeken, Stella, & Di Russo, 2016). In half (i.e., 82) of the experimental stimuli the meanings of X and Y had a literal relationship (Literal, e.g., “that fish is a shark”), while in 82 stimuli the meanings of X and Y had a metaphorical relationship (Metaphors, e.g. “that lawyer is a shark”). To ensure stimuli were all interpreted (and not just retrieved as a whole) we only employed non-lexicalized metaphors. Importantly, the target words were the same in all the stimuli (in the examples above, the word “shark”), while the only difference was in the preceding context. To avoid any carry-over effect stimuli were split into two lists (A e B) each including half of the literal sentences and half of the metaphors, so that each target word appeared once in each list, either in literal or in metaphoric condition. Each participant was administered only with one of the lists that were assigned to each subject in counterbalanced order. To avoid that participants could note that the experiment focus was on metaphors, the final list of stimuli included also 64 fillers, which consisted of further literal sentences, similar to the other stimuli. Fillers were excluded from statistical analyses. Further psycholinguistic details on the stimuli can be found in the Supplementary Materials.

The behavioral task aimed to ascertain that the participants understood the metaphors and, in general, paid attention to the sentences. For this reason, only descriptive statistics on behavioral data were calculated and statistical analyses were conducted only on trials with correct behavioral responses.

2.4. MEG data analysis

The preprocessing of MEG data (see Figure 2) was conducted using Brainstorm (Tadel et al., 2011; version November 2018) in MATLAB 2016b (Mathworks, Inc., Massachusetts, USA), which is available for free download online under the GNU general public license (http://neuroimage.usc.edu/brainstorm). On continuous recordings, we first applied the 3rd gradient noise cancellation. Data were then resampled at 600 Hz and filtered with a notch filter (50 Hz and harmonics at 100, 150, 200, and 250 Hz) and a high-pass filter at 0.1 Hz. The Signal-Space Projection algorithm (SSP) was applied to identify and eliminate cardiac and eye movement artifacts from the recordings. Digital triggers were adjusted offline to match the actual visual stimulus presentation, thereby improving the accuracy of trigger timing.

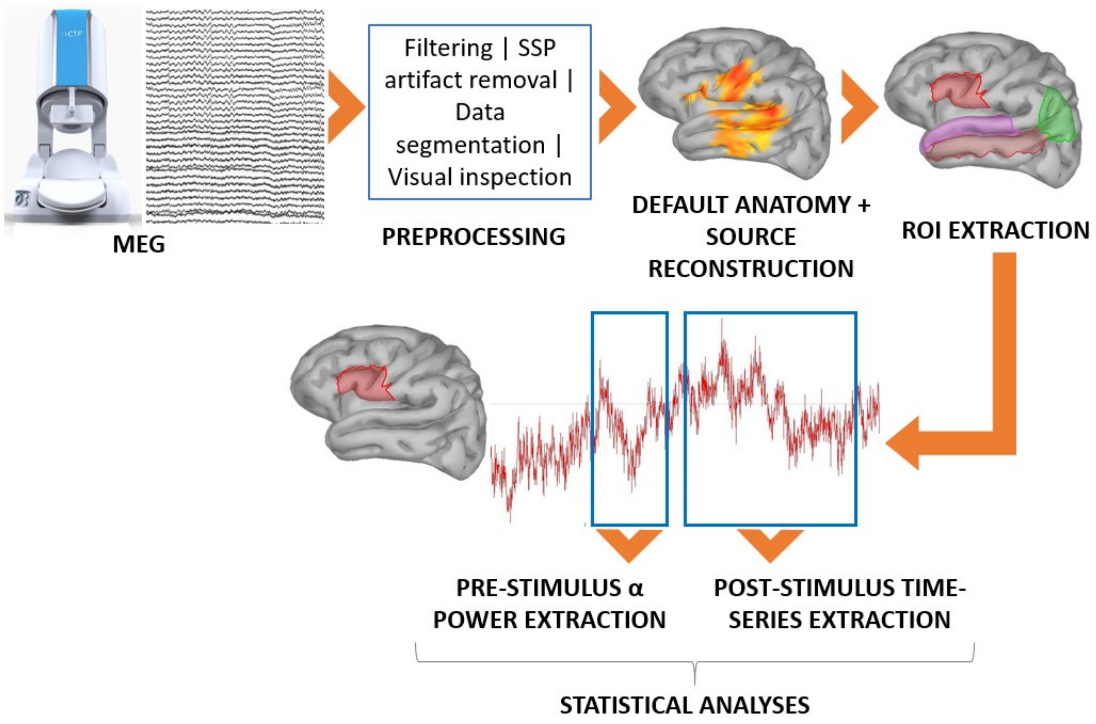


Figure 2. Data collection and analysis steps. The figure shows the main step in data collection and analysis. After data collection, data were cleaned and epoched. Source reconstruction was performed on MEG data and, as a final step before statistical analyses, pre-stimulus single-trial ROI pre-stimulus alpha power, and post-stimulus activation time series were extracted.

Therefore, we extracted epochs that were time-locked to the onset of the target word. The epoch duration ranged from -1500 msec to 1500 msec after the stimulus. During this phase, excessive artifacts were visually inspected and rejected. After trial rejection, each participant had, on average, 36.61 (SD = 2.2) epochs for the *Literal* [range = 29-40] and 36.79 (SD = 2.74) epochs for the *Metaphorical* [range = 30-40], condition. The number of epochs for each condition did not show significant differences on a t-test [t(51.59) = -0.27, p = 0.79].

The MEG forward model was generated using the Boundary Element Method (BEM) based on the default anatomy in Brainstorm. Source reconstruction was performed on the cortex surface using the wMNE (weighted Minimum Norm) algorithm, with Brainstorm's default settings (fixed source orientation, constraining dipoles to be normal to the cortex, depth weighting with Order [0,1] set at 0.5 and Maximal amount set at 10; noise covariance regularization set at 0.1; the regularization parameter 1/λ was determined by setting Signal-To-Noise Ratio to 3). The noise covariance was derived from a 3-minute empty room recording made at the end of each participant's recording session.

Single-trial regions of interest (ROIs) activation time series were reconstructed into 5 cortical ROIs selected from the Destrieux atlas (Destrieux et al., 2010) and dimension-reduced by averaging all signals (i.e. signal from all vertices), within each ROI. As language-related areas, we chose the superior and middle temporal gyri (STG and MTG) and Broca's area. STG and MTG were selected as depicted in the atlas, while Broca's area was obtained by merging the pars opercularis and triangularis (Lorca-Puls et al., 2021). Regarding the TPJs, we took into consideration those portions of this broad region that, according to Doricchi and colleagues (2022), were maximally involved in linguistic computations, since the task of interest was a linguistic one. TPJs were thus obtained by merging two adjacent regions, namely the posterior part of the STG and the angular gyrus (AG; Doricchi et al., 2022), that correspond to cytoarchitectonic areas PGp, PGa (angular gyrus), PF, PFcm, PFm (posterior STG). MNI centroid coordinates for each ROI are reported in Table 1. ROIs included in the analysis are represented in Figure 3.

ROI	x	y	z
Broca	-51	9	12
MTG	-59	-34	-13
STG	-56	-10	-11
lTPJ	-45	-70	37
rTPJ	45	57	39

Table 1. MNI centroid coordinates of the selected ROIs.

We standardized the single-trial, source-reconstructed data through Z-transformation according to the mean value of the 100 msec period preceding the target word, then downsampled to 100 Hz to reduce computational load. Subsequently, the data were exported to R (R Core Team, 2013) using the erpR package (Arcara & Petrova, 2017) for further statistical analyses. Custom code was developed for data handling and plotting.

The preprocessing for the resting-state data was identical to that of the task data. Following the artifact removal through SSP, the continuous recordings were divided into 3-second epochs. Trial rejection was then conducted based on visual inspection, similar to the process applied in the task recordings.

2.5. Single-trial time-frequency analysis

We conducted time-frequency (TF) analyses on source-reconstructed data at the single-trial level. An initial Morlet wavelet (Mother Wavelet in Brainstorm terms) was constructed with a central frequency of 1 Hz and a time resolution of 3 seconds (full-width half-maximum, FWHM). All other wavelet were built as scaled versions of this initial wavelet, spanning from 1 to 45 Hz, with 1 Hz linear frequency steps. We selected this frequency range to allow for additional analyses beyond the scope of this paper, which will not be presented here. From the single-trial time-frequency decomposition, we extracted the average magnitude in the alpha band (8-13 Hz) and exported the single-trial alpha magnitudes to R for statistical analyses. For the resting-state recordings (previously segmented into 3-second epochs for trial rejection purposes), we followed a similar procedure, with the exception that time-resolved alpha-band magnitudes in the single epochs were first averaged in time to obtain a Morlet-based frequency spectrum. Subsequently, these values were averaged across epochs and exported. We opted to use the average resting alpha power as a baseline, as opposed to other inter-trial intervals, because participants were instructed to blink during these intervals to reduce contamination during the experimental trials. As a result, power estimates from this time window would have been affected by such artifacts.

We obtained pre-stimulus power values for each trial by averaging the TF values across the time window between -500 and -10 msec before the target word. We chose this interval length, even if not event-free, to obtain a reliable estimation of alpha power. It is important to note that there is evidence of specific anticipatory processes happening already at the level of the article during sentence reading (DeLong et al., 2005), which might modulate pre-stimulus alpha power. Formally, the eventual influence of the article included in this interval (“a/an”) on alpha power is not controlled, but we overcame such limitation by means of control analyses (reported in the Supplementary Materials, p. 261) demonstrating that none of the experimental items is systematically associated with a particularly lower (or higher) value of alpha. We can therefore rest assured that the uncontrolled properties of the stimuli did not significantly influence the level of pre-stimulus alpha power. Moreover, the words included in the interval are not content words and they are consistent throughout the sentence stimuli, so their effect on sentence comprehension should be constant across all sentences. In addition, the target word is always in agreement with the article, so sentences do not generate syntactic or semantic violations.

By ending the time window 10 ms before the target, we can prevent any undesired effects of temporal smearing caused by time-frequency analyses, like the temporal leakage of stimulus-related oscillatory activity into the pre-stimulus time window, particularly when the stimulus occurs at time 0 or shortly thereafter (Cohen, 2014).

As the absolute single-trial power value might be influenced by irrelevant factors (e.g., individual participant characteristics, signal quality, etc.), we normalized the pre-stimulus power using a baseline division procedure. This procedure involved dividing the power during the task by the average power during the baseline period (in our case, the resting-state activity). The resulting unit of this normalization is a ratio (Cohen, 2014). Values smaller than 1 indicate that during each pre-stimulus period, a given subject had lower alpha power than during the resting state; values larger than 1 indicate that the subject's pre-stimulus power was higher than the resting state power.

Since the power distributions were skewed, we log-transformed the power values for statistical analysis purposes.

2.6. Statistical analyses

2.6.1. Whole-brain and behavioral data analyses

In the context of another paper on the same data (in preparation), we performed a whole-brain Cluster Based Permutation t-test (Maris & Oostenveld, 2007). These analyses were conducted independently from those presented in this paper, but their results confirm that a difference between the activations recorded in the metaphorical and literal condition was present. To reduce the computational burden, data were downsampled at 200 Hz before the cluster-based permutation. Data for all time points and all vertices in the time window from 0 to 1000 ms were then analyzed, setting the cluster alpha to 0.05 and the minimum number of neighbors to 2.

As regards behavioral data, accuracy rates and reaction times were compared by means of paired samples t-tests.

2.6.2. ROI analyses

After exporting the ROIs single trials time-series and pre-stimulus alpha to R, we restricted the time window of interest from -100 to 800 msec. Data from each ROI were analyzed separately, so we ran a total of 5 models (see Figure 3). The models for the TPJs were different from those of the language-related ROIs.

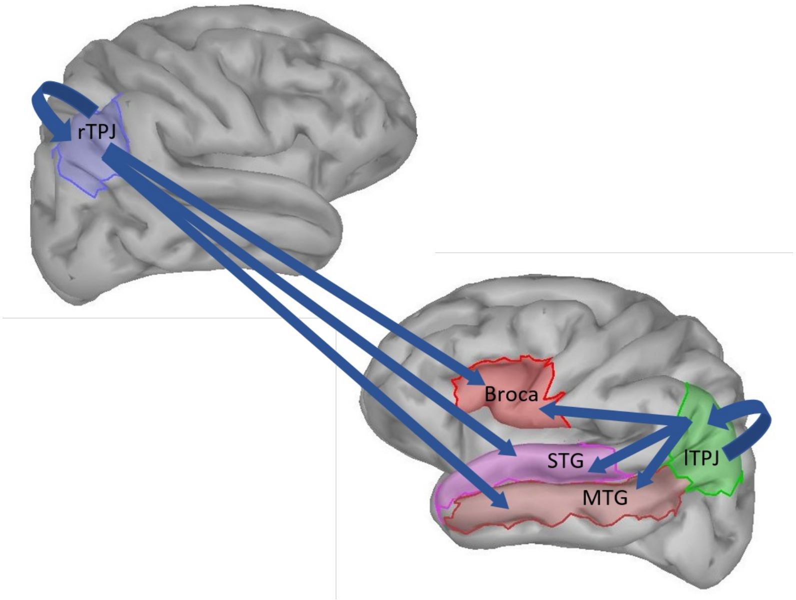


Figure 3. Schematic representation of ROI relationship included in the GAMM models. We extracted pre-stimulus alpha power and post-stimulus activation time course from all the brain areas depicted in the figure. We then constructed the models using the post-stimulus activation as a dependent variable and the pre-stimulus alpha power from the region itself and from the TPJs as main effects (together with other variables, see section 2.6.1.).

In the TPJ models, continuous activation amplitude from the TPJ itself in the whole time window of interest was entered as the dependent variable. Main effects included:

- condition (Metaphor vs. Literal) as factor;
- a non-linear effect of time, depending on condition (this term captures the different changes in the activation over time in the two conditions);
- a non-linear effect of pre-stimulus alpha power from the TPJ itself, capturing the (possibly) non-linear modulation of activation amplitude by different magnitudes of pre-stimulus alpha.
- an interaction term, specifying the non-linear interaction of interest between the continuous variables of time and pre-stimulus alpha power, depending on the condition, was included to capture whether pre-stimulus power modulates, in a possibly nonlinear way, the subsequent activations, in either of the two conditions.

This term corresponds to the interaction of interest that, in GAMMs, is modeled by a tensor smooth function and allows us to answer the question regarding the influence of the prestimulus alpha power on the subsequent activation response within the TPJ.

In the models for the language-related ROIs, additional terms were included, since we wanted to test whether the continuous post-stimulus activation of each ROI (entered as the dependent variable) was better predicted by pre-stimulus alpha power recorded from left or right TPJ, controlling for pre-stimulus alpha power from the same area. Therefore, in addition to the linear main effect of time, the non-linear main effect of time depending on condition, the non-linear main effect of pre-stimulus alpha power from the same ROI, and the interaction term specifying the non-linear interaction of interest between time and pre-stimulus alpha power depending on the condition within the same area (as in the TPJ models), we also added:

- a non-linear effect of pre-stimulus alpha power from rTPJ;
- a non-linear effect of pre-stimulus alpha power from lTPJ;
- an interaction term specifying the non-linear interaction between time and rTPJ pre-stimulus alpha power depending on condition;
- an interaction term specifying the non-linear interaction between time and lTPJ pre-stimulus alpha power depending on condition.

As a consequence, the language ROI models included three interactions of interest, while the TPJ models only included one. This allows us to compare the respective influence of the pre-stimulus alpha from the different locations (rTPJ, lTPJ, or the same ROI) on the post-stimulus activation. Results for the tensor representing the interaction involving the same-area pre-stimulus alpha are reported in the Supplementary Materials.

The R syntax for models with ROI activity on TPJ as dependent variable and for models with Language ROI activity as dependent variable are reported in the Supplementary Materials. In addition to these models, we introduced in all the models a random structure to reduce autocorrelation in the residuals and account for data dependency (van Rij et al., 2019). This included two random factor smooths: one for participants and one for each target word. To identify each trial uniquely, we created a factor variable called "Event," which comprised the combination of participant, condition, and target word. We used this variable as both a random intercept and slope, interacting with time. It's important to note that the target word was included in the random structure solely to model a portion of the random variability between trials and is not a variable of interest in the analysis. Therefore, it will not be discussed further. Additionally, we incorporated an AR1 model to account for autoregressive processes in the data. We estimated the value of rho by examining the model's autocorrelation function (ACF) at Lag 1 and then included it in our model along with a term specifying the starting point of each time series. The nonlinear random structure of the models and the AR1 error model are essential for handling the intrinsic temporal dependency in time series data (van Rij et al., 2019). Since the model's residuals were not normally distributed, we fitted the model with a link function for a scaled-t distribution (Wieling, 2018). To optimize computational time, we set the argument "discrete" to true, enabling more efficient processing. The models showed no convergence issues.

In GAMMs, the effects of the main interactions of interest are mainly interpreted through visual inspection of the tensor functions' plots, which in our case are three-dimensional since they include two continuous predictors: time and pre-stimulus power (for more information on GAMM results visualization, please see Lago et al., 2023). As the sign of results does not help to interpret the data (source activations have not expected signs as ERP components) we rectified the Z-transformed activation time series signal prior to plotting. Any departure from zero can thus be interpreted as “higher activation of the ROI”. 3-D plots are represented with time on the x-axis, pre-stimulus alpha power level on the y-axis, and color-coded activation that provide information on how activation time series are modulated by pre-stimulus power depending on literal and metaphorical conditions. This interaction allows us to visualize the activation time course depending on pre-stimulus alpha levels and conditions, but also to perform subtractions between the metaphorical and literal tensor surfaces as if they were activation time series. We followed this procedure because the time series difference allows us to isolate components of interest that are informative on the effects of the experimental manipulations (Luck, 2005). Moreover, subtracting the metaphorical from the literal tensor surface also allows the identification of significant effects (predicted differences between activation time series) in relation to the pre-stimulus alpha level. These effects were calculated and plotted with the `plot_diff2` function in the package `itsadug` (Van Rij et al., 2020) and are shown in Figure 5.

3. Results

3.1. Behavioral results

In the analysis of the behavioral task, the average Accuracy was 96.40% (SD = 0.18) for Literal sentences and 96.03% (SD = 0.20) for Metaphors. This shows that metaphors were correctly understood. Average RTs were 1311.97 ms (SD = 490.38 ns) for Literal sentences and 1258.97 (SD = 544.48) for Metaphors. There were no significant differences in Accuracy between Metaphors and Literal sentences (t=0.562, *p*=0.574), while the difference in RTs between the two conditions was significant (t=3.089, *p*=0.002), showing that processing literal sentences is probably easier than the processing of metaphors.

3.2. Activation results

3.2.1. Whole-brain results

The cluster-based permutation on whole-brain activation evidenced a significant difference, with higher activation for Metaphors as compared to Literal expressions. This difference (cluster mass = 31850, cluster size = 12033, p = 0.03) was expressed in the left hemisphere with a difference starting at around 620 ms and with activations encompassing the MTG, the STG, and the left Broca's area. No significant differences encompassed the TPJs. Figures depicting this effect can be found in the Supplementary Materials.

3.2.2. ROI results

The percentage of deviance explained by the models ranged between 10.3 and 15.4.

Table 2 reports the significance of the interaction between time, pre-stimulus alpha power, and literal vs. metaphorical condition for left and right TPJ. In these areas, the interaction was not significant. This means that, within the TPJs, pre-stimulus alpha power did not significantly predict the activation in any condition.

Table 2. Beta coefficients and *p* values of interactions between time, power, and condition for the TPJs. There are no significant results.

	lTPJ		rTPJ	
	Beta	p value	Beta	p value
Interaction: time, power, literal condition	0.186	0.654	1.879	0.119
Interaction: time, power, metaphoric condition	3.046	0.080	0.754	0.385

Table 3 reports the significance of the interaction between time, pre-stimulus alpha power, and literal vs metaphorical condition for the language-related ROIs. Within the literal condition, lTPJ pre-stimulus alpha power predicted the post-stimulus activation in all the linguistic ROIs (Broca, MTG, and STG), while rTPJ pre-stimulus power only predicted the post-stimulus activation in the STG. Within the metaphorical condition instead, lTPJ pre-stimulus alpha power did not predict the post-stimulus activation in any of the language-related ROIs, while rTPJ pre-stimulus alpha power only predicted the post-stimulus activation in the STG. This means that pre-stimulus alpha power originating from left and right TPJ significantly predicted the activation in both the literal and metaphorical conditions in different ways across areas (see Figures 4 and 5, only significant effects).

Table 3. Beta coefficients and *p* values of interactions between time, power, and condition for the language-related ROIs. Significant results are marked with *.

	Broca		MTG		STG	
	Beta	p value	Beta	p value	Beta	p value
Interaction: time, lTPJ power, literal condition	4.683	0.031*	6.150	0.014*	2.811	0.006*
Interaction: time, lTPJ power, metaphoric condition	2.156	0.069	0.837	0.632	1.113	0.289
Interaction: time, rTPJ power, literal condition	0.572	0.729	0.658	0.607	3.397	0.021*
Interaction: time, rTPJ power, metaphoric condition	1.319	0.195	1.947	0.059	3.441	0.030

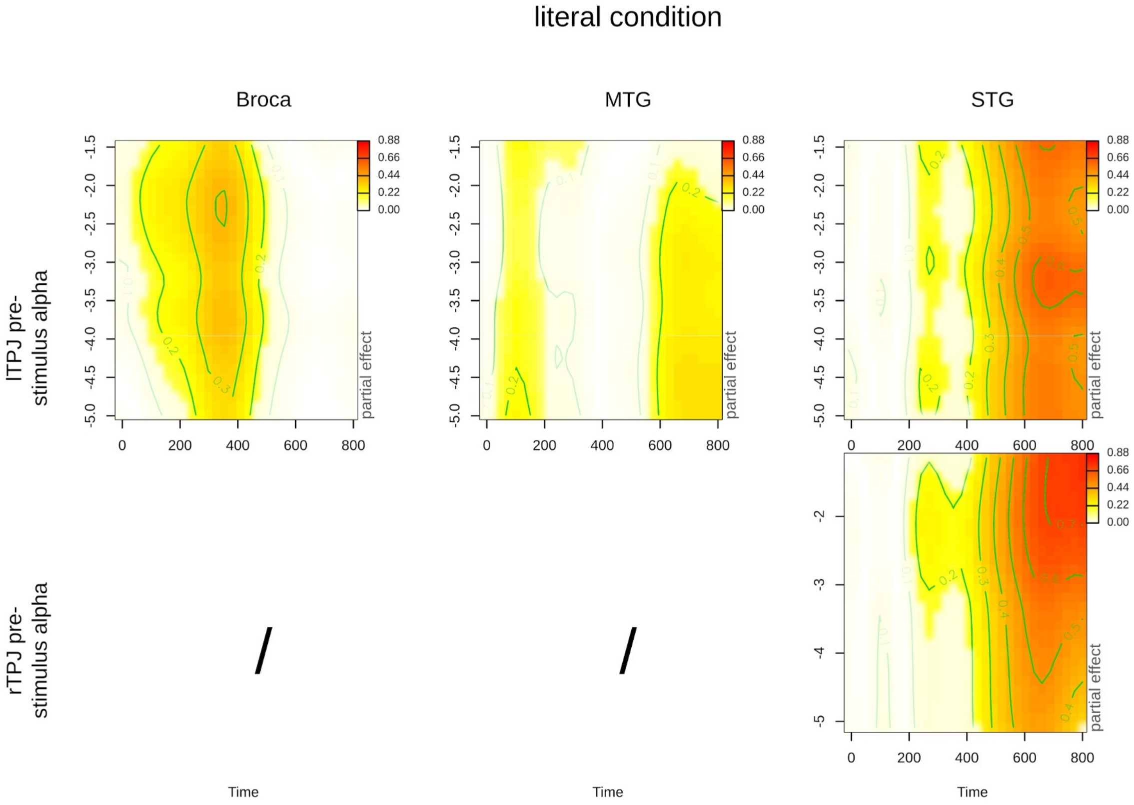


Figure 4. 3-D plots of the main interaction between time and pre-stimulus alpha power for the literal condition, for significant effects. Time is represented on the x-axis, pre-stimulus alpha power level on the y-axis, and activation is color-coded. Colored blots represent plot regions where the interaction is significant, while white areas indicate regions where the confidence intervals (95% CI) around the predicted surface included zero, i.e., the interaction is not significant. Increasingly dark shades indicate greater activation.

In Figures 4 and 5, time is represented on the x-axis, pre-stimulus alpha power level on the y-axis, and activation is color-coded. Colored blots represent plot regions where the interaction is significant, while white areas indicate regions where the confidence intervals (95% CI) around the predicted surface included zero, i.e., the interaction is not significant. Increasingly dark shades indicate greater activation. As we can see from Figure 4, in the literal condition lTPJ pre-stimulus alpha power predicts the post-stimulus activation in Broca's area, the MTG, and STG. In Broca's area, lTPJ pre-stimulus alpha is associated with an early activation (about 200 to 400 ms). In the MTG and STG, it is also associated with activations in later time windows (after 400 ms). Within the STG, post-stimulus activations are also predicted by rTPJ pre-stimulus alpha in late time windows (after 400 ms), similar to the effects observed in association with lTPJ pre-stimulus alpha.

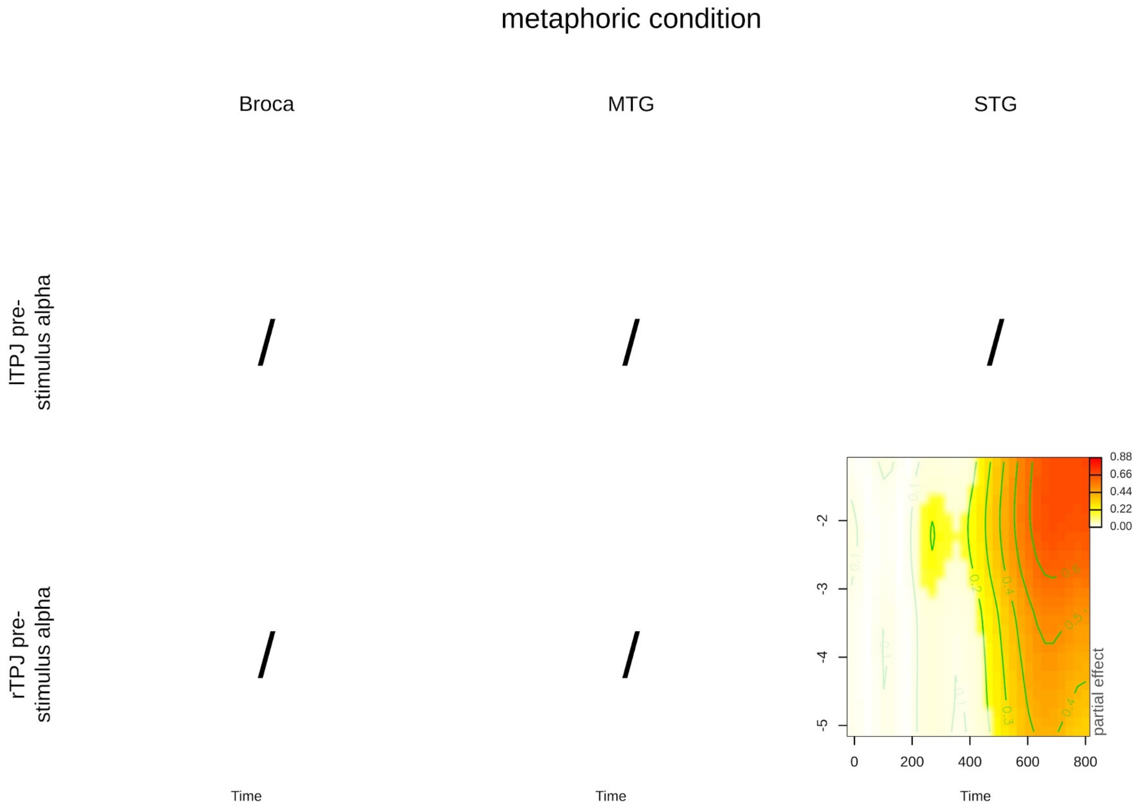


Figure 5. 3-D plots of the main interaction between time and pre-stimulus alpha power for the metaphorical condition, for significant effects. Time is represented on the x-axis, pre-stimulus alpha power level on the y-axis, and activation is color-coded. Colored blots represent plot regions where the interaction is significant, while white areas indicate regions where the confidence intervals (95% CI) around the predicted surface included zero, i.e., the interaction is not significant. Increasingly dark shades indicate greater activation.

Figure 5 instead depicts the interactions for the metaphorical condition. lTPJ pre-stimulus alpha power does not predict the activation in any language-related ROI, while rTPJ pre-stimulus alpha is only associated with late activations in the STG, in a time window resembling the effect in the literal condition.

Since time series differences allow us to isolate components of interest that are informative on the effects of the experimental manipulations (Luck, 2005; see Section 2.6), we subtracted the metaphorical predicted activation from the literal one in each ROI, only for significant interactions. Figure 6 shows, for each ROI, the predicted activation difference between metaphorical and literal conditions depending on the left and right TPJ pre-stimulus alpha level. As in Figures 4 and 5, time is represented on the x-axis, pre-stimulus alpha power level on the y-axis, and predicted activation difference is color-coded. Colored blots represent plot regions where the difference is significant, while white areas indicate regions where the confidence intervals (95% CI) around the predicted activation difference included zero, i.e., the difference is not significant. Increasingly dark shades indicate greater activation differences.

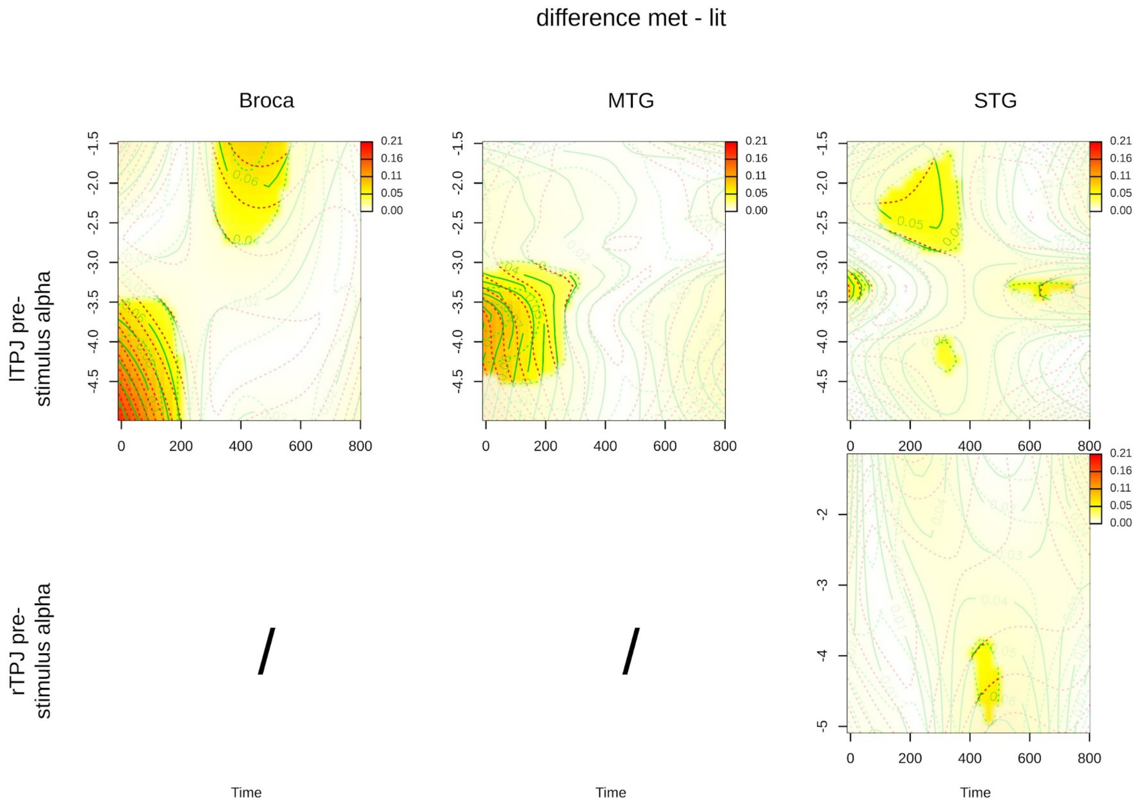


Figure 6. Predicted activation (tensor surface) differences between significant metaphorical and literal conditions for the main interactions of interest (time, alpha band power, and conditions). Time is represented on the x-axis, pre-stimulus alpha power level on the y-axis, and predicted activation the

difference is color-coded. Colored blots represent plot regions where the difference is significant, while white areas indicate regions where the confidence intervals (95% CI) around the predicted activation difference included zero, i.e., the difference is not significant. Increasingly dark shades indicate greater activation differences.

lTPJ pre-stimulus alpha power predicts the activation differences between literal and metaphorical conditions in all the language-related ROIs, albeit in different time segments: lower lTPJ pre-stimulus alpha power is associated with significant activation differences from 0 to 200 msec in Broca's area and the MTG. In Broca's area, higher pre-stimulus alpha is also associated with a significant activation difference around 400 msec. In the STG, differences are more attenuated and distributed, spanning almost the whole time window. Conversely, rTPJ pre-stimulus alpha only predicts a small activation difference around 400 ms in the STG.

To summarize, results show that left and right TPJ pre-stimulus alpha power does *not* predict post-stimulus activation within the TPJs themselves, but it does predict post-stimulus activations in the language-related ROIs in both the literal and metaphoric condition, as well as the differential activation between conditions, in different ways across areas. More specifically, STG post-stimulus activation is predicted by both left and right TPJ pre-stimulus alpha power, while Broca and MTG post-stimulus activations are only predicted by lTPJ pre-stimulus alpha.

4. Discussion

In the present study, we investigated whether the TPJs could be involved in both linguistic prediction generation and testing. In particular, we used Generalized Additive Mixed Models (GAMMs) to test whether the pre-stimulus activity recorded from the TPJs could modulate the post-stimulus activations within the TPJs and in some core linguistic areas. We found that lTPJ pre-stimulus alpha was associated with the post-stimulus activation of all the language-related ROIs under investigation, and it predicted early differences between metaphoric and literal conditions in Broca's area, MTG, and STG. The pattern of results is summarized in Figure 7. In addition, in Broca's area, lTPJ pre-stimulus alpha was associated with significant differences around 400 msec. Here and in the MTG, post-stimulus activation is also predicted by the pre-stimulus alpha power recorded within these same areas (see Supplementary Materials). Importantly, the analytic method used (i.e. GAMM) allowed identifying that the two contributions are separable, hence that the TPJs had some influence in predicting post-stimulus activity of the other areas, despite the fact that the pre-stimulus power within the same area was relevant or not.

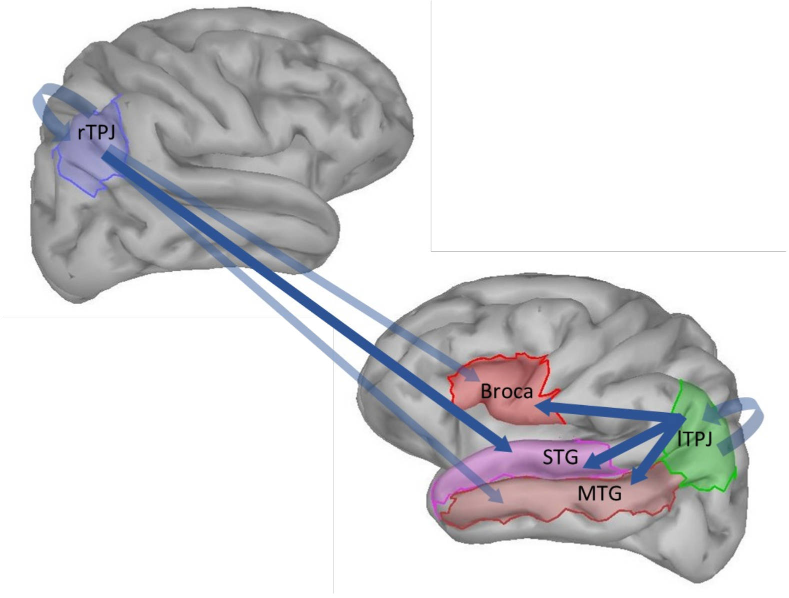


Figure 7. Summary of significant interactions of interest between pre-stimulus alpha power and post-stimulus ROI activations. Solid arrows represent significant interactions, while semitransparent arrows represent the non-significant ones. rTPJ pre-stimulus power significantly predicted the activation difference between metaphorical and literal conditions in the STG. lTPJ pre-stimulus power significantly predicted the activation difference between metaphorical and literal conditions in Broca's area, the MTG, and STG. Each effect is significant for different alpha values and different time points (see Fig. 6).

Conversely, rTPJ only predicted a difference around 400 msec in the STG. What's even more interesting is that within the TPJs, pre-stimulus alpha power did not predict the post-stimulus activation of the areas themselves: in other words, during a metaphor comprehension task, TPJ pre-stimulus alpha power appears to modulate the activation in language-related ROIs, but it does not modulate the activity within the TPJs.

These results indicate that probably, left and right TPJ have different roles in the generation of linguistic predictions, thanks to their functional connections with task-specific ROIs (Masina et al., 2022; Siman-Tov et al., 2019).

4.1. rTPJ

One of the core functions of rTPJ is attentional reorienting (Corbetta et al., 2008; Corbetta & Shulman, 2002), but this region has been proposed to play an ubiquitous role in many cognitive domains (e.g., Masina et al., 2022). Such a pervasive involvement of rTPJ can be explained in terms of predictive processing and, in particular, through the concept of precision, i.e., the reliability of a prediction. In our task, it was the final target word that defined the sentence as literal or metaphoric, therefore we can hypothesize that while reading the sentence, before encountering the target word, participants formed a precise expectation about a literal meaning. As outlined in Section 1.4.1., highly precise expectations about the possible meaning of a sentence could reduce the uncertainty associated with the underlying cause of the sensory input (this has been traditionally interpreted as a higher allocation of attentional resources; Hohwy, 2012). As a consequence, internal representations of the literal sentence meaning should be enhanced, as well as the neural response in the case of a metaphorical target, that disconfirms the prediction of a literal meaning and represents prediction error (Kuperberg et al., 2020). This view is supported by evidence showing that cytoarchitectonic areas of rTPJ, in particular, PGa, PFm, and PF, are sensitive to invalid predictions (invalidly cued trials) in an attentional task. These areas could be specifically recruited when subjects reorient towards internal or external salient stimuli or, in the present case, highly precise hypotheses (Gillebert et al., 2013).

This mechanism might explain the significant effect of low rTPJ pre-stimulus alpha on STG post-stimulus activation visible around 400 msec (Fig. 6, lower right panel). This result is consistent with previous EEG evidence showing that low levels of pre-stimulus alpha, associated with highly precise predictions, modulate the signal in the N400 time window following incongruent target words (Lago et al., 2023).

Evidence for the involvement of rTPJ in figurative language processing has been previously reported (Bambini et al., 2011; Spotorno et al., 2012) and corroborated by the present findings. We hypothesize that rTPJ, in the time interval before the presentation of the final target word, helps enhance predictions relative to the most common (in our case literal) meaning of a sentence, through a mechanism of precision optimization. Linguistic predictions are then tested after the presentation of the target word by more language-specific areas, such as the STG.

Our evidence for the involvement of rTPJ in predictive metaphor processing is however limited since we found that this area only modulated post-stimulus activation in the STG for a very short time. lTPJ, instead, exerts a more extended influence on all the linguistic ROIs under investigation.

4.2. lTPJ

Results showed that lTPJ pre-stimulus alpha power predicted early differences between metaphoric and literal conditions in Broca’s area, MTG, and STG. In addition, in Broca’s area, lTPJ pre-stimulus alpha was associated with significant differences around 400 msec, and in the STG, effects were significant also in a later time window spanning from 500 up until almost the end of the time window (see Figure 6, upper panels).

The fact that pre-stimulus alpha power recorded from the lTPJ has more pervasive effects than those recorded from its right-hemisphere homologous could be a consequence of signal propagation due to the contiguity of lTPJ, MTG, and STG. The STG, the closest region to lTPJ, reports the lowest activation in association with lTPJ pre-stimulus alpha (Fig. 6, upper right panel), while activation in Broca’s area seems to be stronger (Fig. 6, upper left panel). Based on our results and on previous evidence showing that lTPJ integrates global world knowledge and local sentence information (Menenti et al., 2008; Metusalem et al., 2012), we can argue that in the time interval preceding the target word, lTPJ is involved in generating internal models of the sentence based on our prior world knowledge, that have a major role in guiding interpretations in conditions of uncertainty (see Section 1.3.). Internal models' predictions are tested against metaphorical target words with the aid of task-specific areas, namely, Broca, the MTG, and the STG. Interestingly, the predictions that are generated in the lTPJ have a precocious influence on the post-stimulus activation time course in these ROIs: significant effects start right after the target word presentation and last up until 400 msec in the case of Broca's area and of the STG, where significant effects also span from 500 msec almost until the end of the time-window.

This sequence of effects is consistent with previous EEG evidence from other metaphor and figurative language comprehension studies. For example, in their metaphor comprehension task, Schneider et al. (2014) found a significant effect as early as ~200 msec after the processing of the target word. Effects in this time window have been also associated with irony comprehension, another manifestation of figurative language (Regel, Coulson, et al., 2010; Regel, Gunter, et al., 2010), to sentence constraint, and predictability of the target word (Federmeier et al., 2005; Lee et al., 2012; Wlotko & Federmeier, 2007), suggesting that internal models' predictions influence the processing of metaphorical targets starting from very early stages after presentation. Our effects around 400 msec and later are in line with other EEG studies of metaphor comprehension, that have detected N400/late positivity complexes (Bambini et al., 2016; Goldstein et al., 2012; Weiland et al., 2014). Tellingly, modulations around 400 msec are observable in Broca's area and are similar to those found in the STG in association with rTPJ pre-stimulus alpha power, probably reflecting the prediction errors that follow the mismatches between literal predictions and metaphorical target words, similarly to the N400 ERP found in EEG studies. In the STG, these deflections are followed by later modulations that could represent a different dynamic emerging after prediction errors. After the mismatches between literal expectations and metaphorical target words have triggered prediction errors, the cognitive system initially fails to integrate the unexpected word into the internal model, and this failure might cause subsequent attempts to make sense of the sentence through processes like reanalysis, repair, or reinterpretation, that allow correctly inferring the plausible implied meaning of the metaphor. This later dynamic, observed in the STG, is similar to the EEG late positivities found in linguistic violation paradigms (Kuperberg et al., 2020).

It is important to underline that we can draw more general interpretations from these results, without advocating for a specific function to be linked to a specific area (for example, internal model creation to lTPJ). Recent views suggest that specific cognitive functions can arise as a result of the interaction between several brain areas, i.e., brain networks (Pessoa, 2022). Within this perspective, the functional unit to be considered in association to the cognitive function under investigation is in fact the *pattern of interaction* between areas that emerges during a task (in our case, the link between the lTPJ and the STG, Broca and MTG, or between the rTPJ and STG), and such functional unit determines, for example, what we interpreted as precision optimization or internal model creation. However, the role of the single brain areas interacting in a network may be dynamic, therefore changing depending on the task at hand and on the different moments during the same task.

In sum, our results show that the two TPJs are involved in linguistic prediction generation during metaphorical sentence comprehension thanks to their functional connections with language-related areas. Importantly, we showed that the activity of the TPJs seems to precede the one of other areas, supporting the idea that their role is actually related to predictive processes. We suggest that both TPJs could function as predictive hubs, bringing together various cognitive processes and types of information, while closely interacting with other brain regions and networks. As hypothesized by Masina et al. (2022), functional connections between domain-general and task-specific regions are crucial for predictive computations in various cognitive domains, and the specific functions of the TPJs may be determined by the networks activated during ongoing activities, such as the language network in this case. As a result, the context- and network-dependent nature of TPJs facilitates the seamless integration between different brain regions and cognitive processes, enabling context-specific networks to effectively utilize their integrative and contextual-updating capabilities in various behavioral and cognitive scenarios.

It is important to note that left and right TPJ exert a different degree of influence on the post-stimulus activations on the various language-related ROIs. This may depend on their respective role: rTPJ might be more sensitive to precision dynamics that enable the allocation of attentional resources to less uncertain predictions. These precision-related dynamics, reflecting on pre-stimulus alpha power, only modulate the activation in the STG in a short time window around 400 msec, probably reflecting the enhancement of linguistic prediction error. On the other hand, lTPJ would be more involved in the construction of an internal model of the sentence that is tested against the final metaphorical target word, a process that extensively involves all the language-related ROIs under investigation, and causes more extended modulations, that might possibly reflect the reanalysis processes following metaphorical target words and their integration with the sentence internal model, allowing for the correct interpretation of the metaphor.

Limitations. There are some limitations in the present study. The current behavioral task was devised to ensure that the participants paid attention to the stimuli and was easy and without time pressure, therefore such a task does not enable any meaningful analysis of behavioral responses: accuracy was at ceiling and there were no significant differences in reaction times between conditions. For this reason, it is difficult to draw a conclusion on whether pre-stimulus alpha power could influence, besides post-stimulus language ROIs activations, also behavioral performance; or on whether the difficulty of the task could maybe modulate the relationship between pre- and post-stimulus alpha power. Therefore, future studies could employ tasks that allow linking prestimulus alpha not only to neural responses but also to behavioral performance. In addition, we did not directly manipulate the level of sentence constraint, with the consequence that the level of predictions' precision fluctuates spontaneously across stimuli, not allowing to systematically link different levels of alpha power to different levels of constraint and precision. Explicitly addressing the manipulation of sentential constraint could therefore lead to a clearer picture of how more or less precise predictions affect the subsequent stimulus processing. In relation to this point, we have to point out that in language experiments (and the present makes no exception) it is often difficult to rule out the role of uncontrolled properties of the stimuli on the observedeffects. More specifically, our pre-stimulus interval also included the indefinite article, and there is evidence that anticipatory processes might be happening already at this level (DeLong et al., 2005), which might influence the level of pre-stimulus alpha power and, in turn, also the final results. To explore this issue, we ran control analyses based on simulations that can be found in the Supplementary Materials (p. 261). These additional analyses demonstrated that no experimental item is associated with a particular value of alpha, and therefore the effects we report are not due to uncontrolled properties of the stimuli, rather than to random fluctuations. It is also important to note that although we showed a relationship between the activity in both TPJs and language-related areas, the nature of MEG data (as a neurophysiological technique) is purely correlational. Future studies perturbing TPJ activity in tasks and co-registering activity in other areas (e.g. TMS-EEG studies) could further corroborate the interpretations of the results put forward in the present study. In addition, we have to point out that prediction as a continuous and ongoing process, so it it plausible that the TPJs exert a continuous influence on other brain areas throughout the sentence, and that TPJs alpha level could influence the subsequent activation in the language-related areas during a time interval spanning the whole sentence. In the present study instead we took into consideration only specific time intervals at the end of the sentence, since we were interested in studying the difference in activations between metaphors and literal sentences. Future studies could therefore adopt a more complex "sliding window" approach to consider the influence of pre-stimulus alpha on subsequent activation, that would further clarify the complex dynamics of linguistic prediction Finally, an interesting recent perspective (Doricchi et al., 2022) suggests two different and domain-general roles for bilateral TPJ related to match and mismatch. To test this hypothesis a further control would be needed (i.e., neutral cues) that was not available in the current experiment; another strategy could be to perform more fine-grained analyses that circumscribe the investigation of the pre-stimulus dynamics to different areas of the TPJ. For example, Gillebert et al. (2013) found that some cytoarchitectonic areas within rTPJ, in particular, PGa, PFm, and PF, are sensitive to invalid predictions in an attentional task, while another study proposed that neighboring but distinct regions of rTPJ may be selective for Theory of Mind and attention (Scholz et al., 2009). Therefore, future studies could elaborate more on this hypothesis by investigating whether different portions of the TPJs are differently sensitive to (dis)confirmed predictions.

Conclusion. Our findings demonstrate that both TPJs are involved in prediction generation during a metaphor comprehension task, possibly thanks to their functional connections with domain-specific areas. Left and right TPJ are probably involved in linguistic predictive computations with different roles, as they exert a different degree of influence on the post-stimulus activity of language-related brain regions, depending on the area under investigation.

References

Arcara, G. & Petrova, A. erpR: Event-related potentials (ERP) analysis, graphics and utility functions (R version 0.2.0). https://rdrr.io/cran/erpR/ (2017).

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016). Disentangling metaphor from context: An ERP study. *Frontiers in Psychology*, *7*(MAY), 1–14.

Bambini, V., Gentili, C., Ricciardi, E., Bertinetto, P. M., & Pietrini, P. (2011).

Decomposing metaphor processing at the cognitive and neural level through functional magnetic resonance imaging. *Brain Research Bulletin*, *86*(3–4), 203–216. https://doi.org/10.1016/j.brainresbull.2011.07.015

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. https://doi.org/10.1523/JNEUROSCI.3474-13.2014

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, *17*(6), 905–917. https://doi.org/10.1162/0898929054021102

Blanke, O., Ortigue, S., Landis, T., & Seeck, M. (2002). Stimulating illusory own-body perceptions. *Nature*, *419*(September), 269–270.

Bohrn, I., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language—A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia*, *50*, 2669–2683.

Brown, M., & Kuperberg, G. R. (2015). A hierarchical generative framework of language processing: Linking language perception, interpretation, and production abnormalities in schizophrenia. *Frontiers in Human Neuroscience*, *9*(NOV), 1–23. https://doi.org/10.3389/fnhum.2015.00643

Cao, L., Thut, G., & Gross, J. (2017). The role of brain oscillations in predicting self-generated sounds. *NeuroImage*, *147*(November 2016), 895–903. https://doi.org/10.1016/j.neuroimage.2016.11.001

Carter, R. M. K., & Huettel, S. A. (2013). A nexus model of the temporal-parietal junction. *Trends in Cognitive Sciences*, *17*(7), 328–336.

https://doi.org/10.1016/j.tics.2013.05.007

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

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

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: Recollection, familiarity, and novelty. *Journal of Neurophysiology*, *96*(4), 1902–1911. https://doi.org/10.1152/jn.01029.2005

De Ridder, D., Van Laere, K., Dupont, P., Menovsky, T., & Van de Heyning, P. (2007). Visualizing Out-of-Body Experience in the Brain. *New England Journal of Medicine*, *357*(18), 1829–1833. https://doi.org/10.1056/nejmoa070010

Diaz, M. T., & Eppes, A. (2018). Factors influencing right hemisphere engagement during metaphor comprehension. *Frontiers in Psychology*, *9*(MAR), 1–6. https://doi.org/10.3389/fpsyg.2018.00414

Doricchi, F., Lasaponara, S., Pazzaglia, M., & Silvetti, M. (2022). Left and right temporal-parietal junctions (TPJs) as “match/mismatch” hedonic machines: A unifying account of TPJ function. *Physics of Life Reviews*, *42*, 56–92. https://doi.org/10.1016/j.plrev.2022.07.001

Doricchi, F., Macci, E., Silvetti, M., & Macaluso, E. (2010). Neural correlates of the spatial and expectancy components of endogenous and stimulus-driven orienting of attention in the posner task. *Cerebral Cortex*, *20*(7), 1574–1585. https://doi.org/10.1093/cercor/bhp215

Ergenoglu, T., Demiralp, T., Bayraktaroglu, Z., Ergen, M., Beydagi, H., & Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Cognitive Brain Research*, *20*(3), 376–383. https://doi.org/10.1016/j.cogbrainres.2004.03.009

Federmeier, K. D., Mai, H., & Kutas, M. (2005). Both sides get the point: Hemispheric sensitivities to sentential constraint. *Memory & Cognition*, *33*(5), 871–886.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, *4*(December), 1–23. https://doi.org/10.3389/fnhum.2010.00215

Fiebach, C. J., Friederici, A. D., Müller, K., & Von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, *14*(1), 11–23. https://doi.org/10.1162/089892902317205285

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*(JUL), 1–13. https://doi.org/10.3389/fpsyg.2011.00154

Friederici, A. D. (2017). Evolution of the neural language network. *Psychonomic Bulletin and Review*, *24*(1), 41–47. https://doi.org/10.3758/s13423-016-1090-x

Friston, K. (2012). Prediction, perception and agency. *International Journal of Psychophysiology*, *83*(2), 248–252. https://doi.org/10.1016/j.ijpsycho.2011.11.014

Friston, K. (2018). Does predictive coding have a future? *Nature Neuroscience*, *21*(8), 1019–1021. https://doi.org/10.1038/s41593-018-0200-7

Friston, K., Square, Q., & Kilner, J. (2011). Europe PMC Funders Group Action understanding and active inference. *Biological Cybernetics*, *104*(1–2), 137–160. https://doi.org/10.1007/s00422-011-0424-z.Action

Gastaldon, S., Arcara, G., Navarrete, E., & Peressotti, F. (2020). Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. *Cortex*, *133*, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026

Gastaldon, S., Busan, P., Arcara, G., & Peressotti, F. (2023). Inefficient speech-motor control affects predictive speech comprehension: Atypical electrophysiological correlates in stuttering. *Cerebral Cortex*, 1–18. https://doi.org/10.1093/cercor/bhad004

Geng, J. J., & Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: Contextual updating? *Neuroscience and Biobehavioral Reviews*, *37*(10), 2608–2620. https://doi.org/10.1016/j.neubiorev.2013.08.010

Goldstein, A., Arzouan, Y., & Faust, M. (2012). Killing a novel metaphor and reviving a dead one: ERP correlates of metaphor conventionalization. *Brain and Language*, *123*(2), 137–142. https://doi.org/10.1016/j.bandl.2012.09.008

Goodman, N. D., & Frank, M. C. (2016). Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences*, *20*(11), 818–829. https://doi.org/10.1016/j.tics.2016.08.005

Grent-'T-Jong, T., Boehler, C. N., Kenemans, J. L., & Woldorff, M. G. (2011). Differential functional roles of slow-wave and oscillatory-alpha activity in visual sensory cortex during anticipatory visual-spatial attention. *Cerebral Cortex*, *21*(10), 2204–2216. https://doi.org/10.1093/cercor/bhq279

Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in Psychology*, *3*(APR), 1–14. https://doi.org/10.3389/fpsyg.2012.00096

Hughes, C., Cassidy, B. S., Faskowitz, J., Avena-Koenigsberger, A., Sporns, O., & Krendl, A. C. (2019). Age differences in specific neural connections within the Default Mode Network underlie theory of mind. *NeuroImage*, *191*, 269–277. https://doi.org/10.1016/j.neuroimage.2019.02.024

Iemi, L., Busch, N. A., Laudini, A., Haegens, S., Samaha, J., Villringer, A., & Nikulin, V. V. (2019). Multiple mechanisms link prestimulus neural oscillations to sensory responses. *eLife*, *8*(e43620), 1–34. https://doi.org/10.1101/461558

Iemi, L., Chaumon, M., Crouzet, S. M., & Busch, N. A. (2017). Spontaneous neural oscillations bias perception by modulating baseline excitability. *Journal of Neuroscience*, *37*(4), 807–819. https://doi.org/10.1523/JNEUROSCI.1432-16.2016

Indovina, I., & Macaluso, E. (2007). Dissociation of stimulus relevance and saliency factors during shifts of visuospatial attention. *Cerebral Cortex*, *17*(7), 1701–1711. https://doi.org/10.1093/cercor/bhl081

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, *4*(November), 1–8. https://doi.org/10.3389/fnhum.2010.00186

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A Tale of Two Positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, *32*(1), 12–35. https://doi.org/10.1162/jocn_a_01465

Lago, S., Pezzetta, R., Gastaldon, S., Peressotti, F., & Arcara, G. (2023). Trial-by-trial fluctuations of pre-stimulus alpha power predict language ERPs. *Psychophysiology*, 00:e14388. https://doi.org/10.1111/psyp.14388

Lee, C.-Y., Liu, Y.-N., & Tsai, J.-L. (2012). The Time Course of Contextual Effects on Visual Word Recognition. *Frontiers in Psychology*, *3*. https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00285

León-Cabrera, P., Rodríguez-Fornells, A., & Morís, J. (2017). Electrophysiological correlates of semantic anticipation during speech comprehension. *Neuropsychologia*, *99*, 326–334. https://doi.org/10.1016/j.neuropsychologia.2017.02.026

Lorca-Puls, D. L., Gajardo-Vidal, A., Oberhuber, M., Prejawa, S., Hope, T. M. H., Leff, A. P., Green, D. W., & Price, C. J. (2021). Brain regions that support accurate speech production after damage to Broca's area. *Brain Communications*, *3*(4), 1–19. https://doi.org/10.1093/braincomms/fcab230

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Masina, F., Pezzetta, R., Lago, S., Mantini, D., Scarpazza, C., & Arcara, G. (2022). Disconnection from prediction: A systematic review on the role of right temporoparietal junction in aberrant predictive processing. *Neuroscience & Biobehavioral Reviews*, *138*(January), 104713–104713. https://doi.org/10.1016/j.neubiorev.2022.104713

Menenti, L., Petersson, K. M., Scheeringa, R., & Hagoort, P. (2008). When Elephants Fly: Differential Sensitivity of Right and Left Inferior Frontal Gyri to Discourse and World Knowledge. *Journal of Cognitive Neuroscience*, *21*(12), 2358–2368.

Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & Elman, J. L. (2012).

Generalized event knowledge activation during online sentence comprehension. *Journal of Memory and Language*, *66*(4), 545–567. https://doi.org/10.1016/j.jml.2012.01.001

Min, B. K., & Herrmann, C. S. (2007). Prestimulus EEG alpha activity reflects prestimulus top-down processing. *Neuroscience Letters*, *422*(2), 131–135. https://doi.org/10.1016/j.neulet.2007.06.013

Nave, K., Deane, G., Miller, M., & Clark, A. (2020). Wilding the predictive brain. *Wiley Interdisciplinary Reviews: Cognitive Science*, *11*(6), 1–11. https://doi.org/10.1002/wcs.1542

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Prince, S. E., Tsukiura, T., & Cabeza, R. (2007). Distinguishing the neural correlates of episodic memory encoding and semantic memory retrieval. *Psychological Science*, *18*(2), 144–151. https://doi.org/10.1111/j.1467-9280.2007.01864.x

Rapp, A. M., Mutschler, D. E., & Erb, M. (2012). Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. *NeuroImage*, *63*(1), 600–610. https://doi.org/10.1016/j.neuroimage.2012.06.022

Regel, S., Coulson, S., & Gunter, T. C. (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain Research*, *1311*, 121–135. https://doi.org/10.1016/j.brainres.2009.10.077

Regel, S., Gunter, T. C., & Friederici, A. D. (2010). Isn’t It Ironic? An Electrophysiological Exploration of Figurative Language Processing. *Journal of Cognitive Neuroscience*, *23*(2), 277–293.

Rommers, J., Dickson, D. S., Norton, J. J. S., Wlotko, E. W., & Federmeier, K. D. (2017). Alpha and theta band dynamics related to sentential constraint and word expectancy. *Language, Cognition and Neuroscience*, *32*(5), 576–589. https://doi.org/10.1080/23273798.2016.1183799

Ryskin, R., & Nieuwland, M. S. (n.d.). Prediction during language comprehension: What is next? *Trends in Cognitive Sciences*.

Samaha, J., Boutonnet, B., Postle, B. R., & Lupyan, G. (2018). Effects of meaningfulness on perception: Alpha-band oscillations carry perceptual expectations and influence early visual responses. *Scientific Reports*, *8*(1), 1–14. https://doi.org/10.1038/s41598-018-25093-5

Schneider, S., Rapp, A. M., Haeußinger, F. B., Ernst, L. H., Hamm, F., Fallgatter, A. J., & Ehlis, A. C. (2014). Beyond the N400: Complementary access to early neural correlates of novel metaphor comprehension using combined electrophysiological and haemodynamic measurements. *Cortex*, *53*(1), 45–59. https://doi.org/10.1016/j.cortex.2014.01.008

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, *42*, 9–34. https://doi.org/10.1016/j.neubiorev.2014.01.009

Sedley, W., Gander, P. E., Kumar, S., Kovach, C. K., Oya, H., Kawasaki, H., Howard, M. A., & Griffiths, T. D. (2016). Neural signatures of perceptual inference. *eLife*, *5*(MARCH2016), 1–13. https://doi.org/10.7554/eLife.11476

Sherman, M. T., Kanay, R., Seth, A. K., & VanRullen, R. (2016). Rhythmic Influence of Top–Down Perceptual Priors in the Phase of Prestimulus Occipital Alpha Oscillations. *Journal of Cognitive Neuroscience*, *28*(9), 1318-1330-1318–1330.

https://doi.org/10.1162/jocn_a_00973

Siman-Tov, T., Granot, R. Y., Shany, O., Singer, N., Hendler, T., & Gordon, C. R. (2019). Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction. *Neuroscience and Biobehavioral Reviews*, *105*, 262–275. https://doi.org/10.1016/j.neubiorev.2019.08.012

Spotorno, N., Koun, E., Prado, J., Van Der Henst, J. B., & Noveck, I. A. (2012). Neural evidence that utterance-processing entails mentalizing: The case of irony. *NeuroImage*, *63*(1), 25–39. https://doi.org/10.1016/j.neuroimage.2012.06.046

Terporten, R., Schoffelen, J. M., Dai, B., Hagoort, P., & Kösem, A. (2019). The Relation between Alpha/Beta Oscillations and the Encoding of Sentence induced Contextual Information. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-56600-x

Van Den Berg, B., Appelbaum, L. G., Clark, K., Lorist, M. M., & Woldorff, M. G. (2016). Visual search performance is predicted by both prestimulus and poststimulus electrical brain activity. *Scientific Reports*, *6*(May), 1–13. https://doi.org/10.1038/srep37718

van Rij, J., Hendriks, P., van Rijn, H., Baayen, R. H., & Wood, S. N. (2019). Analyzing the Time Course of Pupillometric Data. *Trends in Hearing*, *23*, 1–22. https://doi.org/10.1177/2331216519832483

van Rij J, Wieling M, Baayen R, van Rijn H (2020). “itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs.” R package version 2.4.

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2010). Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience*, *22*(8), 1682–1700. https://doi.org/10.1162/jocn.2009.21293

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *Neuroscientist*, *20*(2), 150–159. https://doi.org/10.1177/1073858413494269

Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, *1464*(1), 242–268. https://doi.org/10.1111/nyas.14321

Wang, L., Hagoort, P., & Jensen, O. (2018). Language Prediction Is Reflected by Coupling between Frontal Gamma and Posterior Alpha Oscillations. *Journal of Cognitive Neuroscience*, *30*(3), 432-447-432–447.

Weiland, H., Bambini, V., & Schumacher, P. B. (2014). The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. *Frontiers in Human Neuroscience*, *8*(AUG), 1–17. https://doi.org/10.3389/fnhum.2014.00583

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, *70*, 86–116. https://doi.org/10.1016/j.wocn.2018.03.002

Wlotko, E. W., & Federmeier, K. D. (2007). Finding the Right Word: Hemispheric Asymmetries in the Use of Sentence Context Information. *Neuropsychologia*, *45*(13), 3001–3014. https://doi.org/10.1016/j.neuropsychologia.2007.05.013

Yonelinas, A. P., Otten, L. J., Shaw, R. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, *25*(11), 3002–3008. https://doi.org/10.1523/JNEUROSCI.5295-04.2005

General Conclusions, Limitations and Future directions

The aim of the present dissertation was to elucidate the role of the bilateral TPJs in neurological patients and healthy individuals, in domain-general cognition and, more specifically, in the language domain.

Study 1, a systematic literature review, investigated the potential role of rTPJ in domain-general cognition under a predictive processing perspective, providing an overview of cognitive impairments in neurological patients as the consequence of structural or functional disconnections or damage to rTPJ. Results of this review confirm the involvement of rTPJ across several tasks and neurological pathologies: via its connections with other brain networks, rTPJ integrates diverse information and updates internal models of the world. Against traditional views, which tend to focus on distinct domains, our hypothesis is that the role of rTPJ can be parsimoniously interpreted as a key hub involved in domain-general predictions, and in case of a direct lesion or a disconnection of rTPJ from other brain areas or networks, as in neurological pathologies, aberrant forms of predictive processing can be observed. This alternative account of rTPJ role in aberrant predictive processing opens different perspectives, stimulating new hypotheses in basic research and clinical contexts.

After establishing preliminary evidence for the role of rTPJ in predictive processing, we aimed to better elucidate the involvement of the bilateral TPJs in prediction generation and testing. To do so, we needed to clarify the more suitable neurophysiological correlates to focus on during the following steps.

Therefore, Study 2 temporarily shifted the focus away from TPJ and demonstrated that spontaneous pre-stimulus alpha power fluctuations taking place on a trial-by-trial basis might be associated with random variations in the ability to precisely predict the upcoming target words and to enhance their relative internal representation. In particular, lower levels of pre-stimulus alpha power should correspond to the ability to hold more precise expectations about the target. As a consequence, predictions about the upcoming targets are enhanced, and generate larger prediction errors (N400s) following incongruent (i.e., unpredicted) targets. After a prediction error is generated, the cognitive system attempts to resolve it through a process of reanalysis (reflected in late posterior positivities/P600s). This process might be associated with changes in precision weights with respect to the precision configuration of the pre-stimulus interval: this evidence suggests that post-stimulus computations such as input reanalysis and prediction error resolution may be dependent on and modulated by the precision level that was driving expectations in the pre-stimulus interval, which can be reflected in varying levels of pre-stimulus alpha power.

These results indicate that investigating precision fluctuations during the pre-stimulus interval might be crucial for a detailed understanding of predictive computations. Notably, attention (traditionally related to alpha modulation) has been considered as an emergent property of the precision optimization mechanism taking place during prediction. In this way, attending to a stimulus means representing and increasing the precision of sensory information (and prediction error) during the inferential process, and the extent to which an individual is attending to a stimulus or prediction can be measured through pre-stimulus alpha oscillations.

Study 3 finally combined the investigations on predictions' precision and on the TPJs put forward in the first two studies, by investigating the neurophysiological dynamics of the TPJs both before and after target stimuli, in functional connection with some core language-related areas during a metaphor comprehension task that entails complex predictive computations (Vespignani et al., 2010). Results show that pre-stimulus alpha power recorded from the left and right TPJ did not modulate the post-stimulus activations within the TPJs themselves, but it exerted different degrees of influence on the post-stimulus activations of the language-related ROIs. Importantly, the activity of the TPJs seems to precede the one on other areas, supporting the idea that their role may be related to prediction generation or maintenance. This suggests that both TPJs could function as predictive hubs (albeit with different roles), bringing together various cognitive processes and types of information, while closely interacting with other brain regions and networks. This finding reinforces the hypothesis, put forward in Study 1, that the TPJs are functionally connected with other domain-specific areas: in particular, rTPJ, being part of the VAN, might coordinate precision optimization in the time interval before the presentation of the final target word, traditionally interpreted as attentional resources allocation (Hohwy, 2012). lTPJ, on the other hand, might have a more domain-specific role in language, by integrating our prior world knowledge and incoming information to generate internal models, which guide interpretations in conditions of uncertainty. Internal models' predictions are subsequently tested against target words with the aid of task-specific areas. This could mean that the specific functions of the TPJs may be determined by the networks activated during ongoing activities, such as the language network in this case (Pessoa, 2022). As a result, the context- and network-dependent nature of TPJs facilitates the seamless integration between different brain regions and cognitive processes, enabling context-specific networks to effectively utilize their integrative and contextual-updating capabilities in various behavioral and cognitive scenarios.

The results of these studies fit into the overarching predictive framework outlined in the General Introduction and contribute to a broader understanding of TPJ's predictive role in domain-specific and domain-general cognition, in healthy participants and neurological patients. Taken together, domain-specific results of studies 2 and 3 on healthy participants can be interpreted by keeping in mind the general conclusions of study 1, which suggests a context- and network-dependent, domain-general role of rTPJ. Study 3 suggests that both left and right TPJ are involved in the generation of complex linguistic predictions, which might then be tested against the actual input with the aid of other language-related areas. This means that the TPJs might not be strictly part of the language network, but they nonetheless modulate the activation level of core linguistic areas during sentence processing, in which they become involved. This result is a great example of the predictive conceptualization put forward in Study 1, i.e., that TPJ has a domain-general role that can be declined in many cognitive domains, depending on the context and on the network it activates.

Predictive theoretical assumptions also propose that prediction is a continuous and ongoing process (León-Cabrera et al., 2017) and therefore, when investigating it, it is crucial to take into account the pre-stimulus dynamics. Study 2 further confirmed this assumption, by demonstrating that such dynamics can exert a great influence on the post-stimulus responses, probably because they carry information on the precision level of the hypothesis that is formulated.

It is also valuable to point out the differences between the methodologies used in the three studies because they enrich the point of view on predictive dynamics. For example, while Study 1 focused on neurological patients, the remaining ones employed healthy participants and had a deeper understanding of the neurophysiological underpinnings of predictive processing, that can be impaired in case of neurological diseases. It would be interesting to deepen the understanding of which mechanisms are affected, and this should be definitively the matter for future studies. Another, yet more stubtle, difference in methodology can be found between study 2 (using EEG) and 3 (using MEG). Study 2 focused on basic ERPs traditionally studied in language research, with the known limitation of a low spatial resolution. This shortcoming was filled by Study 3, which used MEG to localize the effects and give a more complex picture of the spatial and temporal unfolding of linguistic predictive dynamics.

In spite of these methodological differences, or maybe right because of them, the results outlined so far pave the way for the joint investigation of predictions' precision levels, pre-stimulus dynamics, and TPJs activation, which might give a more complete picture of the complex mechanisms of predictive processing.

Limitations and future directions. These studies have some limitations. For example, Study 1 only considered neurological populations and not healthy participants and psychiatric populations. The reason behind this choice can be found in the aim of summarizing findings with a certain clinical relevance, to promote a predictive, more holistic view of neurological diseases. Circumscribing this work to neurological patients opens the field for a new, more parsimonious interpretation of the cognitive impariments resulting from neurological pathologies but also for future investigations on the role of rTPJ in healthy individuals. The study did not investigate the role of lTPJ which may represent another important piece of the puzzle, but focused on rTPJ also because dominant theoretical accounts of the function of the TPJs concentrate mainly on attention. Moreover, long-standing neuropsychological evidence shows that hemispatial neglect is statistically one of the most frequent sequelae after a stroke (Vallar & Calzolari, 2018), and mainly follows lesions to rTPJ (Corbetta & Shulman, 2011). For these reasons, rTPJ has been investigated more extensively with respect to its left-hemisphere counterpart (see General Introduction) and.

On the other hand, Studies 2 and 3 suffer from different limitations, common in neurophysiological studies. One limitation is that behavioral responses were not analyzed because the interest was mainly on the neurophysiological dynamics, therefore the task was designed so that there was no time pressure to respond, thus making response times meaningless. Future studies could employ more difficult tasks that might enable linking pre-stimulus alpha not only to post-stimulus neural responses, but also to behavioral ones. Study 2 also had a more specific limitation, which is intrinsic to most ERP studies on language: it is very difficult to demonstrate that the trial-by-trial alpha variations are actually spontaneous fluctuations and are not instead related to some uncontrolled properties of the stimuli. In short, since all stimuli were not equal, it could be the case that some specific cue-target pairs were associated to lower prestimulus alpha and in turn to enhanced N400 and P600. To explore this issue, additional analyses were performed, that can be found in the

Supplementary Materials (Section 3. in Study 2 Supplementary Materials). These additional analyses demonstrated that the effects we report seem not due to uncontrolled properties of the stimuli. In addition, in Study 3, we did not manipulate the sentence constraint level. This is also an issue that should be addressed in fiture studies, that could systematically vary the level of constraint, to enable investigating whether it could be associated to different levels of pre-stimulus alpha.

In spite of these limitations, our results open new basic research questions to be addressed in future investigations, as outlined in the previous paragraphs, either by experimental task adjustments, or by non-invasive stimulation directed to the TPJs, with the aim of studying how modulations of post-stimulus responses are influenced by manipulations of pre-stimulus activity. Besides basic research, there are open questions also in the field of clinical neuroscience: the bilateral TPJs are closely connected with a number of other areas, and they are part of extended brain networks (e.g., the DAN and DMN), with the functional role of integrating diverse information and updating internal models and expectations. Moreover, as outlined in Study 1, rTPJ lesions are found in many neurological pathologies and cause impairments in a number of cognitive domains, which can be parsimoniously interpreted as a disruption of predictive processes. A deeper understanding of the predictive role of left and right TPJ within task-specific networks paves the way for interpreting a wide range of cognitive impairments resulting from neurological diseases and for developing new treatments aimed at better contrasting the impairment suffered by neurological patients.

Conclusion. The present dissertation thoroughly studied the role of left and right TPJ, both in neurological patients and healthy individuals. As regards neurological populations, our interest was in aberrant forms of predictive processing, while regarding healthy individuals, in the relative unfolding of the neurophysiological correlates of complex predictive computations. In particular, we investigated how pre-stimulus dynamics could influence post-stimulus neural responses and we demonstrated that pre-stimulus alpha power is a candidate correlate for predictions' precision level and, as such, it can modulate classical linguistic ERPs like the N400 and the late posterior positivity/P600. In addition, results demonstrated how both TPJs have a role in elaborating predictions, possibly thanks to their functional connections to domain-specific areas: activity in the alpha band recorded in the TPJs during the time interval preceding the target stimulus can modulate the post-stimulus activation of language-related regions, showing that the processing of task-relevant stimuli is in part dependent to TPJs' state before the stimuli itself, probably indicative of the predictions' precision level. This confirms that the TPJs are crucial predictive hubs and their pre-stimulus state can influence neural responses to target stimuli.

General references

Alamia, A., & VanRullen, R. (2019). Alpha oscillations and travelling waves: Signatures of predictive coding? PLoS Biology, 17(10). https://doi.org/10.1101/464933

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. https://doi.org/10.1523/JNEUROSCI.3474-13.2014

Binder, J. R. (2015). The Wernicke area: Modern evidence and a reinterpretation. Neurology, 85(24), 2170–2175.

Blanke, O., Ortigue, S., Landis, T., & Seeck, M. (2002). Stimulating illusory own-body perceptions. Nature, 419(September), 269–270.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a neurobiologically plausible model of language-related, negative event-related potentials. Frontiers in Psychology, 10(FEB), 1–17. https://doi.org/10.3389/fpsyg.2019.00298

Cabeza, R., Ciaramelli, E., & Moscovitch, M. (2012). Cognitive contributions of the ventral parietal cortex: An integrative theoretical account. Trends in Cognitive

Sciences, 16(6), 338–352. https://doi.org/10.1016/j.tics.2012.04.008

Cao, L., Thut, G., & Gross, J. (2017). The role of brain oscillations in predicting self-generated sounds. NeuroImage, 147(November 2016), 895–903. https://doi.org/10.1016/j.neuroimage.2016.11.001

Carter, R. M. K., & Huettel, S. A. (2013). A nexus model of the temporal-parietal junction. Trends in Cognitive Sciences, 17(7), 328–336. https://doi.org/10.1016/j.tics.2013.05.007

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences, 36(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. Neuron, 58(3), 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews Neuroscience, 3(3), 201–215. https://doi.org/10.1038/nrn755

Corbetta, M., & Shulman, G. L. (2011). Spatial neglect and attention networks (Vol. 34, p. 599). https://doi.org/10.1146/annurev-neuro-061010-113731

Corlett, P. R., Mollick, J. A., & Kober, H. (2022). Meta-analysis of human prediction error for incentives, perception, cognition, and action. Neuropsychofarmacology, 47, 1339–1349. https://doi.org/10.1038/s41386-021-01264-3

Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: Recollection, familiarity, and novelty. Journal of Neurophysiology, 96(4), 1902–1911. https://doi.org/10.1152/jn.01029.2005

De Ridder, D., Van Laere, K., Dupont, P., Menovsky, T., & Van de Heyning, P. (2007). Visualizing Out-of-Body Experience in the Brain. New England Journal of Medicine, 357(18), 1829–1833. https://doi.org/10.1056/nejmoa070010

Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. Neuroscientist, 13(6), 580–593. https://doi.org/10.1177/1073858407304654

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. https://doi.org/10.1038/nn1504

Doricchi, F., Lasaponara, S., Pazzaglia, M., & Silvetti, M. (2022). Left and right temporal-parietal junctions (TPJs) as “match/mismatch” hedonic machines: A unifying account of TPJ function. Physics of Life Reviews, 42, 56–92. https://doi.org/10.1016/j.plrev.2022.07.001

Doricchi, F., Macci, E., Silvetti, M., & Macaluso, E. (2010). Neural correlates of the spatial and expectancy components of endogenous and stimulus-driven orienting of attention in the posner task. Cerebral Cortex, 20(7), 1574–1585. https://doi.org/10.1093/cercor/bhp215

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. Frontiers in Human Neuroscience, 4(December), 1–23. https://doi.org/10.3389/fnhum.2010.00215

Ficco, L., Mancuso, L., Manuello, J., Teneggi, A., Liloia, D., Duca, S., Costa, T., Kovacs, G. Z., & Cauda, F. (2021). Disentangling predictive processing in the brain: A meta-analytic study in favour of a predictive network. Scientific Reports, 11(1), 1–14. https://doi.org/10.1038/s41598-021-95603-5

Fiebach, C. J., Friederici, A. D., Müller, K., & Von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. Journal of Cognitive Neuroscience, 14(1), 11–23. https://doi.org/10.1162/089892902317205285

Friston, K. (2012). Prediction, perception and agency. International Journal of Psychophysiology, 83(2), 248–252. https://doi.org/10.1016/j.ijpsycho.2011.11.014

Friston, K. (2018). Does predictive coding have a future? Nature Neuroscience, 21(8), 1019–1021. https://doi.org/10.1038/s41593-018-0200-7

Friston, K. (2019). Waves of prediction. PLoS Biology, 17(10), 1–7.

https://doi.org/10.1371/journal.pbio.3000426

Gastaldon, S., Arcara, G., Navarrete, E., & Peressotti, F. (2020). Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. Cortex, 133, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026

Geng, J. J., & Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: Contextual updating? Neuroscience and Biobehavioral Reviews, 37(10), 2608–2620. https://doi.org/10.1016/j.neubiorev.2013.08.010

Gillebert, C. R., Mantini, D., Peeters, R., Dupont, P., & Vandenberghe, R. (2013). Cytoarchitectonic mapping of attentional selection and reorienting in parietal cortex. NeuroImage, 67, 257–272. http://dx.doi.org/10.1016/j.neuroimage.2012.11.026

Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. Frontiers in Psychology, 3(APR), 1–14. https://doi.org/10.3389/fpsyg.2012.00096

Hughes, C., Cassidy, B. S., Faskowitz, J., Avena-Koenigsberger, A., Sporns, O., & Krendl, A. C. (2019). Age differences in specific neural connections within the Default Mode Network underlie theory of mind. NeuroImage, 191, 269–277. https://doi.org/10.1016/j.neuroimage.2019.02.024

Igelström, K. M., & Graziano, M. S. A. (2017). The inferior parietal lobule and temporoparietal junction: A network perspective. Neuropsychologia, 105(September 2016), 70–83. https://doi.org/10.1016/j.neuropsychologia.2017.01.001

Indovina, I., & Macaluso, E. (2007). Dissociation of stimulus relevance and saliency factors during shifts of visuospatial attention. Cerebral Cortex, 17(7), 1701–1711. https://doi.org/10.1093/cercor/bhl081

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. Frontiers in Human Neuroscience, 4(November), 1–8. https://doi.org/10.3389/fnhum.2010.00186

Jensen, O., Pan, Y., Frisson, S., & Wang, L. (2021). An oscillatory pipelining mechanism supporting previewing during visual exploration and reading. Trends in

Cognitive Sciences, 25(12), 1033–1044. https://doi.org/10.1016/j.tics.2021.08.008

Jiang, J., Summerfield, C., & Egner, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. Journal of Neuroscience, 33(47), 18438–18447. https://doi.org/10.1523/JNEUROSCI.3308-13.2013

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. Brain Research Reviews, 53(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Kutas, M., & Hillyard, S. A. (1984). Brain Potentials During Reading Reflect Word Expectancy and Semantic Association. Nature, 307(5947), 161–163.

Leckey, M., & Federmeier, K. D. (2020). The P3b and P600(s): Positive contributions to language comprehension. Psychophysiology, 57(7), 1–15. https://doi.org/10.1111/psyp.13351

León-Cabrera, P., Rodríguez-Fornells, A., & Morís, J. (2017). Electrophysiological correlates of semantic anticipation during speech comprehension. Neuropsychologia, 99, 326–334. https://doi.org/10.1016/j.neuropsychologia.2017.02.026

Masina, F., Pezzetta, R., Lago, S., Mantini, D., Scarpazza, C., & Arcara, G. (2022). Disconnection from prediction: A systematic review on the role of right temporoparietal junction in aberrant predictive processing. Neuroscience & Biobehavioral Reviews, 138(January), 104713–104713. https://doi.org/10.1016/j.neubiorev.2022.104713

Menenti, L., Petersson, K. M., Scheeringa, R., & Hagoort, P. (2008). When Elephants Fly: Differential Sensitivity of Right and Left Inferior Frontal Gyri to Discourse and World Knowledge. Journal of Cognitive Neuroscience, 21(12), 2358–2368.

Nicenboim, B., Vasishth, S., & Rösler, F. (2020). Are words pre-activated probabilistically during sentence comprehension? Evidence from new data and a bayesian random-effects meta-analysis using publicly available data. Neuropsychologia, 142(107427).

https://doi.org/10.1016/j.neuropsychologia.2020.107427

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., Ferguson, H. J., Fu, X., Heyselaar, E., Huettig, F., Husband, E. M., Ito, A., Kazanina, N., Kogan, V., Kohút, Z., Kulakova, E., Mézière, D., Politzer-Ahles, S., Rousselet, G., … Von Grebmer Zu Wolfsthurn, S. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a largescale study using brain potentials. Philosophical Transactions of the Royal Society B: Biological Sciences, 375(20180522). https://doi.org/10.1098/rstb.2018.0522

Pan, Y., Popov, T., Frisson, S., & Jensen, O. (2023). Saccades are locked to the phase of alpha oscillations during natural reading. PLoS Biology, 21(1), e3001968. https://doi.org/10.1371/journal.pbio.3001968

Pessoa, L. (2022). The Entangled Brain. The MIT Press.

Poldrack, R. A., & Yarkoni, T. (2016). From brain maps to cognitive ontologies: Informatics and the search for mental structure. Annual Review of Psychology, 67, 587–612. https://doi.org/10.1146/annurev-psych-122414-033729

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. Clinical Neurophysiology, 118(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Prince, S. E., Tsukiura, T., & Cabeza, R. (2007). Distinguishing the neural correlates of episodic memory encoding and semantic memory retrieval. Psychological Science, 18(2), 144–151. https://doi.org/10.1111/j.1467-9280.2007.01864.x

Sani, I. (2023). The missing link among the functions of the temporo-parietal junction Comment on “Left and right temporal-parietal junctions (TPJs) as «match/mismatch» hedonic machines: A unifying account of TPJ function” by Doricchi et al. Physics of Life Reviews, 44, 9–12. https://doi.org/10.1016/j.plrev.2022.11.006

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. Neuroscience and Biobehavioral Reviews, 42, 9–34.

https://doi.org/10.1016/j.neubiorev.2014.01.009

Schurz, M., Tholen, M. G., Perner, J., Mars, R. B., & Sallet, J. (2017). Specifying the brain anatomy underlying temporo-parietal junction activations for theory of mind: A review using probabilistic atlases from different imaging modalities. Human Brain Mapping, 38(9), 4788–4805. https://doi.org/10.1002/hbm.23675

Sedley, W., Gander, P. E., Kumar, S., Kovach, C. K., Oya, H., Kawasaki, H., Howard, M. A., & Griffiths, T. D. (2016). Neural signatures of perceptual inference. eLife, 5(MARCH2016), 1–13. https://doi.org/10.7554/eLife.11476

Sherman, M. T., Kanay, R., Seth, A. K., & VanRullen, R. (2016). Rhythmic Influence of Top–Down Perceptual Priors in the Phase of Prestimulus Occipital Alpha Oscillations. Journal of Cognitive Neuroscience, 28(9), 1318-1330-1318–1330. https://doi.org/10.1162/jocn_a_00973

Siman-tov, T., Granot, R. Y., Shany, O., Singer, N., Hendler, T., Brain, S., Aviv, T., Medical, S., & Aviv, T. (2019). Is there a prediction network? Meta-analytic evidence for a cortical- subcortical network likely subserving prediction. Neuroscience and Biobehavioral Reviews, 105(August), 262–275. https://doi.org/10.1016/j.neubiorev.2019.08.012

Vallar, G., & Calzolari, E. (2018). Unilateral spatial neglect after posterior parietal damage (1a ed., Vol. 151, p. 312). Elsevier B.V. https://doi.org/10.1016/B978-0-444-63622-5.00014-0

Van De Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwill, D. J. (2010). Monitoring in language perception: Mild and strong conflicts elicit different ERP patterns. Journal of Cognitive Neuroscience, 22(1), 67–82. https://doi.org/10.1162/jocn.2008.21170

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2010). Predictive mechanisms in idiom comprehension. Journal of Cognitive Neuroscience, 22(8), 1682–1700. https://doi.org/10.1162/jocn.2009.21293

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and ventral attention systems:

Distinct neural circuits but collaborative roles. Neuroscientist, 20(2), 150–159. https://doi.org/10.1177/1073858413494269

Walsh, K. S., McGovern, D. P., Clark, A., & O’Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. Annals of the New York Academy of Sciences, 1464(1), 242–268. https://doi.org/10.1111/nyas.14321

Yonelinas, A. P., Otten, L. J., Shaw, R. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. Journal of Neuroscience, 25(11), 3002–3008. https://doi.org/10.1523/JNEUROSCI.5295-04.2005

Supplementary Materials

Study 1

1. Databases and algorithms

This supplementary material contains all details about the full search strategy for all databases considered in the systematic review. Last date of search: 23rd October 2020.

1.1 Pubmed

((right [Title/Abstract]) AND (TPJ[Title/Abstract] OR temporal parietal junction[Title/Abstract] OR temporoparietal junction[Title/Abstract] OR temporal-parietal junction[Title/Abstract] OR temporo-parietal junction[Title/Abstract] OR temporo-parietal region[Title/Abstract])) OR (rTPJ[Title/Abstract]) AND (lesion* [Title/Abstract] OR dementia[MeSh Terms] OR mild cognitive impairment[Title/Abstract] OR Parkinson disease[MeSh Terms] OR multiple sclerosis[MeSH Terms] OR stroke[MeSH Terms] OR brain damage[MeSH Terms] OR dementia[Title/Abstract] OR Alzheimer[Title/Abstract] OR Parkinson[Title/Abstract] OR Huntington[Title/Abstract] OR multiple sclerosis[Title/Abstract] OR neurological[Title/Abstract])

Total records = 169

1.2 Proquest

https://search.proquest.com/advanced

AB("right") AND (AB("TPJ") OR AB("temporal parietal junction") OR AB("temporoparietal junction") OR AB("temporal-parietal junction") OR AB("temporo-parietal junction") OR AB("temporo-parietal region") OR AB(rTPJ)) AND (AB(lesion*) OR MESH.EXACT("Dementia") OR MESH.EXACT(“Cognitive Dysfunction”) OR MESH.EXACT("Parkinson Disease") OR MESH.EXACT("Multiple Sclerosis") OR MESH.EXACT("Stroke") OR AB("brain damage") OR AB("dementia") OR AB("alzheimer") OR AB("parkinson") OR AB("huntington") OR AB("multiple sclerosis") OR AB("neurological"))

Total records = 151 (peer reviewed and English filters)

1.3 Embase

('right':ab,ti AND ('tpj':ab,ti OR 'temporal parietal junction':ab,ti OR 'temporoparietal junction':ab,ti OR 'temporal-parietal junction':ab,ti OR 'temporo-parietal junction':ab,ti OR 'temporo-parietal region':ab,ti) OR 'rtpj':ab,ti) AND ('lesion*':ab,ti OR 'dementia'/mj OR 'mild cognitive impairment'/mj OR 'parkinson disease'/mj OR 'multiple sclerosis'/mj OR 'cerebrovascular accident'/mj OR 'brain damage'/mj OR 'dementia':ab,ti OR 'alzheimer':ab,ti OR 'parkinson':ab,ti OR 'huntington':ab,ti OR 'multiple sclerosis':ab,ti OR 'neurological':ab,ti)

Total records = 219

2. Tables

Table S1. The items of the modified version of the Newcastle-Ottawa Scale (NOS, Wells et al., 2011). In this systematic review, we used an adapted version for cross-sectional studies (Patra et al., 2015) that we customized according to the aims of our systematic search.

	Question 1	**Question 2**	**Question 3**	**Question 4**
Selection (max 4 points)	*Is the case definition adequate?* a) yes, with independent validation (1 point) b) yes, e.g. record linkage or based on self reports c) no description	*Representativeness of the cases:* a) consecutive or obviously representative series of cases (1 point) b) potential for selection biases or not stated	*Selection of Controls:* a) community controls (1 point) b) hospital controls c) no description	*Definition of Controls:* a) no history of disease (endpoint; 1 point) b) no description of source
Comparability (max 2 points)	*Comparability of cases and controls on the basis of the design or analysis:* a) study controls for the most important factor (age; 1 point) b) study controls for any additional factor (education; 1 point)			

Outcome (max 3 points)	*Ascertainment of variable of interest:* a) Validated tool/measure (1 point) b) Non-validated tool/measure, but the tool/measure is available or described and is replicable (1 point) c) No description of the measurement tool	*Same method of ascertainment for cases and controls:* a) yes (1 point) b) no	*Statistical analysis:* a) The statistical analysis used to analyze the variable of interest is clearly described and appropriate, and confidence intervals, p values, or effect sizes are reported (1 point) b) The statistical analysis is not appropriate, not described or incomplete (trends reported as significance and/or no correction for multiple comparison (without clear a-priori).

Table S2. Quality assessment using the Newcastle-Ottawa scale (NOS).

Authors	**Title**	**Year**	**Pathology**	**Design**	**Selection**					**Comparability**			**Outcome**				**Overall**	**Risk of bias**
Agosta et al.,	The Pivotal Role of the Right Parietal Lobe in Temporal Attention	2017	Acquired brain injury	Case-control	a	a	c	a	3	a		1	a	a	a	3	**7**	**low**
Boccia et al.,	Topological and hodological aspects of body representation in right brain damaged patients	2020	Acquired brain injury	Case-control	a	a	a	a	4	a	b	2	b	a	a	3	**9**	**low**

Cohen-Zimerman et al.,	The neural basis for mental state attribution: A voxel-based lesion mapping study	2020	Acquired brain injury	Case-control	a	b	a	a	3	a	b	2	a	a	a	3	**8**	**low**
Kaski et al.,	Temporoparietal encoding of space and time during vestibular-guided orientation	2016	Acquired brain injury	Case-control	a	b	c	a	2	a		1	a	a	a	3	**6**	**medium**
Leigh et al.,	Acute lesions that impair affective empathy	2013	Acquired brain injury	Case-control	a	a	a	a	4	a		1	a	a	a	3	**8**	**low**
Mandonnet et al.,	A network-level approach of cognitive flexibility impairment after surgery of a right temporo-parietal glioma	2017	Acquired brain injury	Single case	a	b	c	b	1			0	a	b	b	1	**2**	**high**
Martinaud et al.,	Ownership illusions in patients with body delusions: Different neural profiles of visual capture and disownership	2017	Acquired brain injury	Case series	a	a	c	b	2			0	a	b	a	2	**4**	**medium**
Monai et al.,	Multiple Network Disconnection in Anosognosia for Hemiplegia	2020	Acquired brain injury	Case series	a	b	c	b	1			0	b	b	a	2	**3**	**high**
Pedrazzini and Ptak,	Damage to the right temporoparietal junction, but not lateral prefrontal or insular cortex, amplifies the role of goal-directed attention	2019	Acquired brain injury	Case-control	a	a	c	b	2	a		1	b	a	a	3	**6**	**medium**
Shomstein et al.,	Top-down and bottom-up attentional guidance: investigating the role of the dorsal and ventral parietal cortices	2010	Acquired brain injury	Case-control	a	b	c	b	1	a	b	2	b	a	b	2	**5**	**medium**
Singh and Knight,	Effects of posterior association cortex lesions on brain potentials preceding self-initiated movements	1993	Acquired brain injury	Case-control	a	b	a	a	3	a		1	b	a	a	3	**7**	**low**
Starkstein et al.,	Anosognosia in patients with cerebrovascular lesions. A study of causative factors	1992	Acquired brain injury	Case series	a	a	c	b	2			0	b	b	a	2	**4**	**medium**
Wawrzyniak et al.,	The neuronal network involved in self-attribution of an artificial	2018	Acquired brain injury	Case series	a	a	c	b	2			0	a	b	a	2	**4**	**medium**

	hand: A lesion network-symptom-mapping study																
Chechlacz et al.,	Neuronal substrates of Corsi Block span: Lesion symptom mapping analyses in relation to attentional competition and spatial bias	2014	Acquired brain injury (neglect)	Case-control	a	b	a	a	3		0	a	b	a	2	**5**	**medium**
Committeri et al.,	Where did you "left" Piazza del Popolo? At your right temporo-parietal junction.	2015	Acquired brain injury (neglect)	Case series	a	a	c	b	2		0	a	b	a	2	**4**	**medium**
Demeurisse et al.,	Pathogenesis of subcortical visuo-spatial neglect. A HMPAO SPECT study	1997	Acquired brain injury (neglect)	Case-control (controls only for behavioral normative data)	a	b	c	b	1		0	a	a	a	3	**4**	**medium**
Dressing et al.,	The correlation between apraxia and neglect in the right hemisphere: A voxel-based lesion-symptom mapping study in 138 acute stroke patients	2020	Acquired brain injury (neglect)	Case-control (controls only for behavioral normative data)	a	a	c	b	2		0	a	a	a	3	**5**	**medium**
Golay et al.,	Cortical and subcortical anatomy of chronic spatial neglect following vascular damage	2008	Acquired brain injury (neglect)	Case series	a	b	c	b	1		0	a	a	a	3	**4**	**medium**
Hattori et al.,	Structural connectivity in spatial attention network: reconstruction from left hemispatial neglect	2018	Acquired brain injury (neglect)	Case-control	a	a	b	a	3	a	1	a	a	b	2	**6**	**medium**
Karnath et al.,	The cortical substrate of visual extinction	2003	Acquired brain injury (neglect)	Case-control	a	a	c	a	3		0	b	a	a	3	**6**	**medium**
Kaufman et al.,	Multiperturbation analysis of distributed neural networks: the case of spatial neglect	2009	Acquired brain injury (neglect)	Case series	a	b	c	b	1		0	a	b	a	2	**3**	**high**
Lee et al.,	Neglect dyslexia: frequency, association with other hemispatial neglects, and lesion localization	2009	Acquired brain injury (neglect)	Case series	a	a	c	b	2		0	a	b	a	2	**4**	**medium**
Pedrazzini and Ptak,	The neuroanatomy of spatial awareness: a large-scale region-of-	2020	Acquired brain injury (neglect)	Case series	a	a	c	b	2		0	a	b	a	2	**4**	**medium**

	interest and voxel-based anatomical study																
Pedrazzini et al.,	A neuroanatomical model of space-based and object-centered processing in spatial neglect	2017	Acquired brain injury (neglect)	Case-control	a	a	c	b	2			0	a	b	a	2	**4**	**medium**
Ptak and Schnider,	The attention network of the human brain: relating structural damage associated with spatial neglect to functional imaging correlates of spatial attention	2011	Acquired brain injury (neglect)	Case-control	a	a	c	b	2	a		1	b	a	a	3	**6**	**medium**
Rousseaux et al.,	Anatomical and psychometric relationships of behavioral neglect in daily living	2015	Acquired brain injury (neglect)	Case series	a	b	c	b	1			0	a	b	a	2	**3**	**high**
Thiebaut de Schotten et al.,	Damage to white matter pathways in subacute and chronic spatial neglect: a group study and 2 single-case studies with complete virtual "in vivo" tractography dissection.	2014	Acquired brain injury (neglect)	Case series	a	b	c	b	1			0	a	a	a	3	**4**	**medium**
Ticini et al.,	The role of temporo-parietal cortex in subcortical visual extinction	2010	Acquired brain injury (neglect)	Case series	a	a	c	b	2			0	a	b	a	2	**4**	**medium**
Toba et al.,	Game theoretical mapping of causal interactions underlying visuo-spatial attention in the human brain based on stroke lesions	2017	Acquired brain injury (neglect)	Case series	a	b	c	b	1			0	b	b	a	2	**3**	**high**
Baez et al.,	Brain structural correlates of executive and social cognition profiles in behavioral variant frontotemporal dementia and elderly bipolar disorder	2019	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	c	a	3	a		1	b	a	a	3	**7**	**low**
Blanc et al.,	Cortical Thickness in Dementia with Lewy Bodies and Alzheimer's Disease: A Comparison of Prodromal and Dementia Stages	2015	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	a	a	4		b	1	b	a	a	3	**8**	**low**
De Marco et	Age and hippocampal	2019	Neurodegenerativ	Case-	a	a	c	b	2	a	b	2	b	a	a	3	**7**	**low**

al.,	volume predict distinct parts of default mode network activity		e disease (dementia or cognitive impairment)	control														
Diez et al.,	Corticolimbic fast-tracking: enhanced multimodal integration in functional neurological disorder	2019	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	a	a	4	a		1	b	a	a	3	**8**	**low**
Eslinger et al.,	Apathy in frontotemporal dementia: behavioral and neuroimaging correlates	2012	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	c	a	3	a	b	2	a	a	a	3	**8**	**low**
Kang et al.,	Effects of Lewy body disease and Alzheimer disease on brain atrophy and cognitive dysfunction	2019	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	c	a	3	a	b	2	b	a	a	3	**8**	**low**
Luks et al.,	Atrophy in two attention networks is associated with performance on a Flanker task in neurodegenerative disease	2010	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	c	a	3			0	b	a	a	3	**6**	**medium**
Qian et al.,	Functional-structural degeneration in dorsal and ventral attention systems for Alzheimer's disease, amnestic mild cognitive impairment	2015	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	c	b	2	a	b	2	b	a	a	3	**7**	**low**
Rohrer et al.,	Alzheimer's pathology in primary progressive aphasia	2012	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	c	c	2			0	b	a	a	3	**5**	**medium**
Sorg et al.,	Asymmetric loss of parietal activity causes spatial bias in prodromal and mild Alzheimer's disease	2012	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	b	a	c	b	1	a	b	2	b	a	a	3	**6**	**medium**
Yamashita et al.,	Functional connectivity change between posterior cingulate cortex and ventral attention network relates to the impairment of orientation for time in Alzheimer's disease patients	2019	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	c	b	2			0	b	a	a	3	**5**	**medium**
Zamboni et al.,	Anosognosia for behavioral disturbances	2010	Neurodegenerative disease	Case-control	a	a	c	a	3	a	b	2	b	a	a	3	**8**	**low**

	in frontotemporal dementia and corticobasal syndrome: A voxel-based morphometry study		(dementia or cognitive impairment)															
Zou et al.,	3.0 T MRI arterial spin labeling and magnetic resonance spectroscopy technology in the application of Alzheimer's disease	2014	Neurodegenerative disease (dementia or cognitive impairment)	Case-control	a	a	c	a	3	a	b	2	b	a	a	3	**8**	**low**
Hanganu et al.,	Depressive symptoms in Parkinson's disease correlate with cortical atrophy over time	2017	Neurodegenerative disease (Parkinson)	Longitudinal	a	a	c	b	2			0	b	b	a	2	**4**	**medium**
Pickut et al.,	Mindfulness based intervention in Parkinson's disease leads to structural brain changes on MRI: a randomized controlled longitudinal trial	2013	Neurodegenerative disease (Parkinson)	Longitudinal	a	a	c	b	2			0	b	b	a	2	**4**	**medium**
Beauchamp et al.,	Electrocorticography Links Human Temporoparietal Junction to Visual Perception	2012	Other pathologies	Case series	b	b	c	c	0			0	b	b	a	2	**2**	**high**
Jiang et al.,	Altered attention networks and DMN in refractory epilepsy: A resting-state functional and causal connectivity study	2018	Other pathologies	Case-control	a	a	c	a	3	a	b	2	b	a	a	3	**8**	**low**
Lu et al.,	Visual rehabilitation training alters attentional networks in hemianopia: An fMRI study	2018	Other pathologies	Longitudinal	a	a	c	b	2			0	b	b	b	1	**3**	**high**
Peterson et al.,	Right Temporoparietal Junction Transcranial Magnetic Stimulation in the Treatment of Psychogenic Nonepileptic Seizures: A Case Series	2018	Other pathologies	Case series	b	b	c	b	0			0	b	b	a	2	**2**	**high**
Zhang et al.,	Transcutaneous auricular vagus nerve stimulation at 1 Hz modulates locus coeruleus activity and resting state functional connectivity in patients with migraine: An fMRI	2019	Other pathologies	Crossover	a	a	c	b	2			0	b	a	a	3	**5**	**medium**

	study																	
Bosma et al.,	Dynamic pain connectome functional connectivity and oscillations reflect multiple sclerosis pain	2018	White matter disease	Case-control	a	a	a	a	4	a		1	b	a	a	3	**8**	**low**
Carotenuto et al.,	Pragmatic abilities in multiple sclerosis: The contribution of the temporoparietal junction	2018	White matter disease	Case series	a	b	c	b	1			0	a	b	a	2	**4**	**medium**
Huang et al.,	White matter lesion loads associated with dynamic functional connectivity within attention network in patients with relapsing-remitting multiple sclerosis	2019	White matter disease	Case-control	a	a	a	a	4	a		1	b	a	a	3	**7**	**low**
Kim et al.,	Neuropathic pain and pain interference are linked to alpha-band slowing and reduced beta-band	2019	White matter disease	Case-control	a	a	c	a	3	a		1	b	a	a	3	**6**	**medium**
Peng et al.,	Density abnormalities in normal-appearing gray matter in the middle-aged brain with white matter hyperintense lesions: a DARTEL-enhanced voxel-based morphometry study	2016	White matter disease	Case-control	a	a	a	a	4	a	b	2	b	a	a	3	**8**	**low**

Table S3. In this table, two kinds of information are specified: (1) “Alteration location”, which includes information about brain structures involved in each study and, if specified, more precise details about what the authors intended with “rTPJ” (e.g., by differentiating between the angular and supramarginal gyrus); (2) “Type of alteration”, which specifies whether rTPJ was directly damaged, disconnected, hypoperfused, and so on.

Authors	**Title**	**Year**	**Alteration location**	**Type of alteration**

Agosta et al.,	The Pivotal Role of the Right Parietal Lobe in Temporal Attention	2017	rTPJ	Direct lesion
Boccia et al.,	Topological and hodological aspects of body representation in right brain damaged patients	2020	Body structural representation: right STG, insula, SMG and TPJ, Rolandic operculum, IFG, and posterior arcuate segment	Direct lesion and disconnection
Cohen-Zimerman et al.,	The neural basis for mental state attribution: A voxel-based lesion mapping study	2020	Right DLPFC, right MFG, left IPL, left ILF, right SLF	Disconnection
Kaski et al.,	Temporoparietal encoding of space and time during vestibular-guided orientation	2016	rTPJ	Direct lesion
Leigh et al.,	Acute lesions that impair affective empathy	2013	Right PFC, OFG, ACC, anterior insula, temporal pole (including TPJ), and amygdala	No alteration
Mandonnet et al.,	A network-level approach of cognitive flexibility impairment after surgery of a right temporo-parietal glioma	2017	rTPJ	Direct lesion
Martinaud et al.,	Ownership illusions in patients with body delusions: Different neural profiles of visual capture and disownership	2017	Disturbed sensation of limb ownership: rTPJ, SMG, MFG	Direct lesion
Monai et al.,	Multiple Network Disconnection in Anosognosia for Hemiplegia	2020	rTPJ, right insula, right lateral and medial prefrontal cortex, SLF III, arcuate fasciculus, fronto-insular, frontal inferior longitudinal, and frontal aslant.	Direct lesion and disconnection
Pedrazzini and Ptak,	Damage to the right temporoparietal junction, but not lateral prefrontal or insular cortex, amplifies the role of goal-directed attention	2019	Right anterior angular gyrus, posterior part of SLF, arcuate fasciculus, fronto-occipital fasciculus	Disconnection

Shomstein et al.,	Top-down and bottom-up attentional guidance: investigating the role of the dorsal and ventral parietal cortices	2010	TPJ, SPL	Direct lesion
Singh and Knight,	Effects of posterior association cortex lesions on brain potentials preceding self-initiated movements	1993	Anterior Brodmann areas 39 and 40 (angular and SMG) and area 7 (SPL, IPL, precuneus)	Direct lesion
Starkstein et al.,	Anosognosia in patients with cerebrovascular lesions. A study of causative factors	1992	Thalamus, rTPJ	Direct lesion
Wawrzyniak et al.,	The neuronal network involved in self-attribution of an artificial hand: A lesion network-symptom-mapping study	2018	TPJ (SMG), right IFG, and right anterior insula	Direct lesion
Chechlacz et al.,	Neuronal substrates of Corsi Block span: Lesion symptom mapping analyses in relation to attentional competition and spatial bias	2014	rTPJ, IPL, MTG, MOG, SLF, arcuate fasciculus, IFOF, and ILF	Direct lesion and disconnection
Committeri et al.,	Where did you "left" Piazza del Popolo? At your right temporo-parietal junction.	2015	Posterior rTPJ (angular gyrus and SMG)	Direct lesion
Demeurisse et al.,	Pathogenesis of subcortical visuo-spatial neglect. A HMPAO SPECT study	1997	Lenticular nucleus, thalamus, rTPJ	Direct lesion
Dressing et al.,	The correlation between apraxia and neglect in the right hemisphere: A voxel-based lesion-symptom mapping study in 138 acute stroke patients	2020	Imitation of meaningless postures: MTG, TPJ, STG and sulcus, angular gyrus, parieto-occipital sulcus, secondary sensorimotor cortex and (peri-)insular regions	Direct lesion
Golay et al.,	Cortical and subcortical anatomy of chronic spatial neglect following vascular damage	2008	rTPJ and underlying white matter, SMG, posterior STG, and insula	Direct lesion

Hattori et al.,	Structural connectivity in spatial attention network: reconstruction from left hemispatial neglect	2018	SLF III/Arcuate fasciculus and the extreme capsule/inferior fronto-occipital fasciculus	Disconnection
Karnath et al.,	The cortical substrate of visual extinction	2003	rTPJ, including the ventral part of the IPL, the caudal parts of the STG and MTG	Direct lesion
Kaufman et al.,	Multiperturbation analysis of distributed neural networks: the case of spatial neglect	2009	Supramarginal and angular gyri of the IPL, the SPL, the anterior part of the rTPJ connecting the STG and SMG, and the thalamus	Direct lesion
Lee et al.,	Neglect dyslexia: frequency, association with other hemispatial neglects, and lesion localization	2009	The STG, MTG, IPL, posterior insular cortex, and the lingual and fusiform gyri	Direct lesion
Pedrazzini and Ptak,	The neuroanatomy of spatial awareness: a large-scale region-of-interest and voxel-based anatomical study	2020	Supramarginal part of the rTPJ and SLF-mediated frontoparietal connections (SLF I, SLF II, and SLF III)	Direct lesion and disconnection
Pedrazzini et al.,	A neuroanatomical model of space-based and object-centered processing in spatial neglect	2017	Anterior part of rTPJ	Direct lesion
Ptak and Schnider,	The attention network of the human brain: relating structural damage associated with spatial neglect to functional imaging correlates of spatial attention	2011	IPS, dorsal premotor cortex, and rTPJ	Direct lesion
Rousseaux et al.,	Anatomical and psychometric relationships of behavioral neglect in daily living	2015	Neglect in daily living was associated with lesions of the posterior part of the STG and extending to rTPJ, temporo-occipital junction and subcortical white matter (including the SLF).	Direct lesion and disconnection

Thiebaut de Schotten et al.,	Damage to white matter pathways in subacute and chronic spatial neglect: a group study and 2 single-case studies with complete virtual "in vivo" tractography dissection	2014	The SLF II	Disconnection
Ticini et al.,	The role of temporo-parietal cortex in subcortical visual extinction	2010	rTPJ and neighboring clusters: the posterior part of STG, the posterior part of MTG, angular gyrus, and SMG. In addition, the IFC	Direct lesion
Toba et al.,	Game theoretical mapping of causal interactions underlying visuo-spatial attention in the human brain based on stroke lesions	2017	IPS, rTPJ, IFG, and IOG	Direct lesion
Baez et al.,	Brain structural correlates of executive and social cognition profiles in behavioral variant frontotemporal dementia and elderly bipolar disorder	2019	Atrophy of rTPJ and superior temporal pole, as well as the amygdala, the hippocampus, the parahippocampal gyrus, the putamen, the insula, and the precuneus, was associated with Theory of Mind impairments	Direct lesion
Blanc et al.,	Cortical Thickness in Dementia with Lewy Bodies and Alzheimer's Disease: A Comparison of Prodromal and Dementia Stages	2015	Cortical thinning was found in the rTPJ, insula, and cingulate, orbitofrontal, and lateral occipital cortices in dementia with Lewy bodies	Direct lesion
De Marco et al.,	Age and hippocampal volume predict distinct parts of default mode network activity	2019	The disconnection between the default-mode network and the anterior portion of rTPJ is related to hippocampal volumes	Disconnection
Diez et al.,	Corticolimbic fast-tracking: enhanced multimodal integration in functional neurological disorder	2019	Altered functional connectivity from the left anterior insula to the right anterior insula and TPJ	Disconnection
Eslinger et al.,	Apathy in frontotemporal dementia: behavioral and neuroimaging correlates	2012	Atrophy in the right caudate (encompassing the ventral striatum), rTPJ, right ITG, right MTG, and left	Direct lesion

			frontal operculum-anterior insula region	
Kang et al.,	Effects of Lewy body disease and Alzheimer disease on brain atrophy and cognitive dysfunction	2019	Lewy body disease and Alzheimer's disease share the degeneration of TPJ and parietal cortices	Direct lesion
Luks et al.,	Atrophy in two attention networks is associated with performance on a Flanker task in neurodegenerative disease	2010	rTPJ, VLPFC, and DLPFC atrophy	Direct lesion
Qian et al.,	Functional-structural degeneration in dorsal and ventral attention systems for Alzheimer's disease, amnestic mild cognitive impairment	2015	Alzheimer's disease is associated with impaired functional connectivity in several areas belonging to attention networks: right OSG/MFG, right IPL, angular gyrus, and SMG around TPJ	Alteration of functional connectivity
Rohrer et al.,	Alzheimer's pathology in primary progressive aphasia	2012	Patients suffering from severe forms of logopenic/phonological aphasia show cortical thinning in the left posterior superior temporal, inferior parietal, medial temporal, posterior cingulate, left anterior temporal, and frontal cortices. In addition, in the right hemisphere, TPJ, posterior cingulate, and medial temporal lobe are involved	Direct lesion
Sorg et al.,	Asymmetric loss of parietal activity causes spatial bias in prodromal and mild Alzheimer's disease	2012	Hypometabolism in IPL and TPJ	Hypometabolism
Yamashita et al.,	Functional connectivity change between posterior cingulate cortex and ventral attention network relates to the impairment of orientation for time in Alzheimer's disease patients	2019	Disconnection between PCC and right MTG (adjacent to rTPJ) is related to orientation for time impairment in Alzheimer's disease patients	Disconnection
Zamboni et al.,	Anosognosia for behavioral disturbances in frontotemporal dementia and	2010	Right STS (close to rTPJ)	Direct lesion

	corticobasal syndrome: A voxel-based morphometry study			
Zou et al.,	3.0 T MRI arterial spin labeling and magnetic resonance spectroscopy technology in the application of Alzheimer's disease	2014	TPJ and frontal, temporal, parietal, and hippocampal regions, bilaterally	Hypoperfusion
Hanganu et al.,	Depressive symptoms in Parkinson's disease correlate with cortical atrophy over time	2017	rTPJ, right occipital medial region, right DLPFC, right posterior cingulate region, left MTG, and left SMA	Direct lesion
Pickut et al.,	Mindfulness based intervention in Parkinson's disease leads to structural brain changes on MRI: a randomized controlled longitudinal trial	2013	Whole brain analysis showed increased gray matter in the left and right caudate nucleus, the left occipital lobe at the lingual gyrus and cuneus, the left thalamus, and bilaterally in TPJ	No alteration
Beauchamp et al.,	Electrocorticography Links Human Temporoparietal Junction to Visual Perception	2012	rTPJ	No alteration
Jiang et al.,	Altered attention networks and DMN in refractory epilepsy: A resting-state functional and causal connectivity study	2018	Ventral attention network: reduction of functional connectivity in VPFC and TPJ, bilaterally; Default-mode network: reduction of functional connectivity in bilateral MPFC; increase of functional connectivity in bilateral precuneus	Alteration of functional connectivity
Lu et al.,	Visual rehabilitation training alters attentional networks in hemianopia: An fMRI study	2018	Increase of functional connectivity between the rTPJ to the insula and ACC	Alteration of functional connectivity
Peterson et al.,	Right Temporoparietal Junction Transcranial Magnetic Stimulation in the Treatment of Psychogenic Nonepileptic Seizures: A Case Series	2018	TPJ (Brodmann area 39)	No alteration
Zhang et al.,	Transcutaneous auricular vagus nerve stimulation at 1 Hz modulates locus coeruleus activity and resting state functional connectivity in patients with migraine: An fMRI study	2019	Transcutaneous auricular vagus nerve stimulation increases connectivity among locus coeruleus, rTPJ, and left	Alteration of functional connectivity

			secondary somatosensory cortex	
Bosma et al.,	Dynamic pain connectome functional connectivity and oscillations reflect multiple sclerosis pain	2018	Default-mode network (rTPJ)	Alteration of functional connectivity
Carotenuto et al.,	Pragmatic abilities in multiple sclerosis: The contribution of the temporoparietal junction	2018	TPJ and paracingulate cortex	Alteration of functional connectivity
Huang et al.,	White matter lesion loads associated with dynamic functional connectivity within attention network in patients with relapsing-remitting multiple sclerosis	2019	Decreased functional connectivity in dorsal attention network (left IPS-right IPS, left IPS-right FEF, left FEF-right IPS) and in ventral attention network (rTPJ-right VFC), along with increased functional connectivity between dorsal attention network and ventral attention network (left IPS-right VFC, left IPS-rTPJ)	Alteration of functional connectivity
Kim et al.,	Neuropathic pain and pain interference are linked to alpha-band slowing and reduced beta-band	2019	Thalamus, posterior insula, and rTPJ	Alteration of spectral power
Peng et al.,	Density abnormalities in normal-appearing gray matter in the middle-aged brain with white matter hyperintense lesions: a DARTEL-enhanced voxel-based morphometry study	2016	Decreased gray matter in the left MFG, bilateral ACC, left and right premotor cortex, and left and right MCC. Increased gray matter in the bilateral cerebellum, left MTG, rTPJ, left and right PFC, and left IPL. Relationship between white matter hyperintense lesion volume and the decreased gray matter density in left MFG, bilateral ACC, right MCC, and right premotor cortex. White matter hyperintense lesion volume negatively correlated with increased gray matter in the rTPJ, left and right PFC, and right PFC	Direct lesion

List of acronyms (Brain structures)

ACC: anterior cingulate cortex; DLPFC: dorsolateral prefrontal cortex; FEF: frontal eye field; IFC: inferior frontal cortex; IFG: inferior frontal gyrus; IFOF: inferior fronto-occipital fasciculus; ILF: inferior longitudinal fasciculus; IOG: inferior occipital gyrus; IPL: inferior parietal lobule; IPS: intraparietal sulcus; ITG: inferior temporal gyrus; MCC: middle cingulate cortex; MFG: middle frontal gyrus; MOG: middle occipital gyrus; MPFC: medial prefrontal cortex; MTG: middle temporal gyrus; OFG: orbitofrontal gyrus; OSG: orbital superior gyrus; PCC: posterior cingulate cortex; PFC: prefrontal cortex; PPC: posterior parietal cortex; SLF: superior longitudinal fasciculus; SMA: supplementary motor area; SMG: supramarginal gyrus; SPL: superior parietal lobule; STG: superior temporal gyrus; STS: superior temporal sulcus; TPJ: temporoparietal junction; VFC: ventral frontal cortex; VLPFC: ventrolateral prefrontal cortex; VPFC: ventral prefrontal cortex.

3. A predictive processing account of neurological disorders

In this supplementary part of the review, we will suggest complementary explanations for deficits related to damage or disconnection of rTPJ considering all of them through the lens of predictive processing. This may provide insights to interpret these deficits as the consequence of the disruption of predictive processing, rather than traditional domain-specific deficits.

3.1 Attention

Within the predictive processing framework precision estimation is functionally analogous to attention (Feldman and Friston, 2010; Hohwy, 2013). Indeed, when individuals pay attention to a stimulus, an object, or a context, the reliability of the error signal is potentially higher than the opposite scenario when attention is low. Thus, the more attention is paid, the more the gathered information is supposedly accurate. If a certain prediction error is assigned to a high degree of reliability, the impact (or the weight) of prediction error on the internal model updating will be higher. In other words, the predictive processing assigns to attention a novel functional meaning, and rTPJ, which is a hub of the VAN and strictly engaged in attention (Corbetta and Shulman, 2002), would represent a pivotal brain structure within the predictive processing network (Siman-Tov et al., 2019).

Emphasis should be laid on spatial attention since its relationship with rTPJ has been steadily found. Spontaneously reorienting attention is necessary to pick up stimuli in an ever-changing environment and neglect patients offer a paradigm in which patients overlook the contralesional side of space, typically the left one. A relatively large number of authors consider neglect as a dysfunction of predictive processing. At a computational level, Parr and Friston (2018) suggested that neglect may be the consequence of abnormal prior beliefs. This hypothesis outlines several scenarios, for example, a patient with neglect may generate abnormal sensorimotor hypotheses (i.e., abnormal *priors*), leading to report low occurrence probability of events in the contralesional hemifield, with the consequence of reducing the tendency to shift attention towards the left side of space. According to another interpretation, the impairment of spatial attention in neglect is thought to be due to a disruption of the VAN, critical for precision estimates and, in turn, the building up of an environmental model during the sampling of left-sided stimuli, rather than the non-impaired sampling of right-sided stimuli (Dietz et al., 2021). As a result, patients with neglect may have difficulties in updating the representation of probabilistic environmental contingencies.

Recently, Doricchi and colleagues (2021) tested the intriguing hypothesis that patients with neglect suffer from impairment to predict probabilistic regularities of sensory events. The authors investigated the electrophysiological markers that were evoked by auditory irregularities, namely the mismatch negativity and the P3, respectively considered as markers of pre-attentive/lower-order prediction errors and higher-order prediction errors. Findings showed that patients had predictive processing impairments both at lower- and higher-order prediction errors, resulting in difficulties to update the representation of probabilistic contingencies in the environment. If we consider these predictions as the basis for model updating and awareness, this result may explain the lack of conscious attention of the left side of space in patients with neglect (Garrido and Deouell, 2021).

3.2 Awareness and Social cognition

rTPJ is also engaged in social cognition and ToM tasks (Bardi et al., 2017; Young et al., 2010), namely tasks that require reasoning about others' mental/emotional states. An important aspect of ToM tasks is that they activate the DMN, which is also associated with internally directed processes such as memory and introspection (Buckner and Carroll, 2007).

The DMN encompasses rTPJ and is anticorrelated with the VAN (Fox et al., 2005) since the latter is active during externally directed cognitive processes.

In sum, rTPJ is engaged in ToM tasks in association with the DMN, and in attention tasks in association with the VAN. Although apparently counterintuitive, this pattern of activation might reflect processes that are active in both kinds of tasks. Decety and Lamm (2007) hypothesized that successful performance in ToM tasks is based on the comparison between internal models about what the others will do/think/feel, which is an internally directed process, with the actual perceptual evidence. This comparison requires the ability to shift between internally and externally directed processes. Corbetta and co-workers (2008) agreed with this view, by hypothesizing that rTPJ might be involved in both attention and ToM because the VAN - thanks to the reorienting function of rTPJ - can shift attention from an internal to an external perspective. In a predictive view of cognition, comparing internal models or predictions with the actual external information is key to optimizing performance by minimizing prediction errors. Carter and Huettel (2013) further expanded this idea by hypothesizing that rTPJ, being active during several tasks - both internally and externally directed - acts as a hub where several cognitive processes converge, with repercussions to decision-making in social context. We therefore posit that rTPJ could be responsible for enabling internal models to predict complex social environments and, in general, to promote social cognition. Such a convergence of different cognitive processes in rTPJ and its involvement in reorienting attention to salient stimuli implies that rTPJ has an integrative function. This not only involves social cognition, but also corporeal awareness, and sense of agency (Decety and Lamm, 2007). Indeed, the disruption or experimental manipulation of rTPJ was found to cause sensations of mismatch between one’s own body and the external environment, like out-of-body experiences. This has been explained by advocating rTPJ as an integration hub between visuospatial and proprioceptive information (Blanke and Arzy, 2005), but results of the current review suggest that its role goes beyond simple sensory integration and is not limited to attention, social cognition, and bodily awareness. The rTPJ role is instead related to the integration and comparison of internal models of task context with the actual external information, across a wide range of cognitive and behavioral tasks.

The idea of the brain as a predictive machine is intriguing but how it can generate, test and update hypotheses regarding a complex and ever-changing environment is still unclear. Some authors agree that motion can be exploited to test hypotheses (i.e., prior beliefs) about proprioceptive and visual sensations. This hypothesis, within the active inference framework, suggests that an individual uses movement in search of sensory feedback. As Fotopoulou (2014) suggested, anosognosia for hemiplegia seems a condition accompanied by an impairment of active inference together with disconnection of the likelihood mapping between motor control and the related intero- and exteroceptive outcomes. Computationally, anosognosia for hemiplegia might reflect aberrant predictive processing, specifically a dynamic imbalance between prior beliefs, sensory feedback, and prediction errors (Fotopoulou, 2014). Experts are still divided on whether to consider anosognosia for hemiplegia as a secondary consequence of sensory deficits or because of impaired movement monitoring. According to the first hypothesis, these patients would be unaware of their paralysis because they are unable to update prior beliefs regarding their motor capabilities. Somehow, they would have difficulties accessing sensory feedback. However, a further plausible explanation is that anosognosia for hemiplegia involves the inability to gather the discrepancy between predicted and actual sensory feedback (Kirsch et al., 2021).

3.3 Motor functions and Apathy

According to the predictive processing framework, motor control occurs through the integration between sensory information and predictions emerging from internal models of action. As already mentioned, this integration is functionally dependent on the level of precision of sensory information and precision of predictions. Within this framework, a variety of motor disturbances can be interpreted as impairment in predictive processing.

In general, in the case of a highly uncertain environment (e.g., a dark room), individuals' movements rely more on their *priors* (e.g., expectations about the light switch position in the room) than sensory evidence (Körding and Wolpert, 2004). Furthermore, in normal aging there is greater precision of predictions and less of sensory evidence rather than in neurodegenerative disorders, such as in PD, in which motor deficits may result from an over-confidence in sensory evidence together with an under-confidence in sensorimotor predictions, leading to poverty of initiating movements (Wolpe et al., 2018). In these patients, the strong adherence to the senses is confirmed in a study showing how movement performance was reduced in case of the reduction of visual feedback (thus visual information) with respect to healthy individuals who do not present this performance decay (Klockgether and Dichgans, 1994). Interestingly, peripheral vibrotactile stimulation seems to reduce motor symptoms in PD patients by increasing the precision of *priors* and, at the same time, reducing the precision from sensory evidence (Macerollo et al., 2018).

More related to higher-order motivational deficits, leading to a difficulty in initiating goal-directed behaviors, Kocagoncu and colleagues (2020) suggested that in apathy the reduction of movements would depend on the integration of rewards and costs in predictions. Patients with apathy might expect low rewards or high costs from their actions with the consequence of reducing goal-directed movements. Apathy may entail uncertainty in predicting the outcome of one's own actions and, as reported by Jakobs and co-workers (2009), rTPJ appears to have a central role in predictive motor processing, especially in conditions characterized by higher uncertainty where rTPJ activation is higher. As the authors stated, this suggests that rTPJ engagement may be linked to increased updating of prior beliefs in case of higher uncertainty.

3.4 Executive functions

Executive dysfunctions are commonly observed in most brain disorders (Godefroy et al., 2018; Moreira et al., 2017), leading patients to be more disinhibited, impulsive, and prone to perseverative errors. The Wisconsin Card Sorting Test has been one of the most used neuropsychological tools to evaluate executive functions, in particular cognitive flexibility. Typically, patients with poor cognitive flexibility show the tendency to perseverate with inadequate responses despite a shift in the context/rule that would require different responses. In a recent study, Barcelo (2020) provided convincing evidence supporting the interpretation of patients' outcomes at this task as the result of predictive impairments. Crucially, similarly to Doricchi and colleagues (2021), the author showed how P3-like responses would index high- and low-level expectation updating offering biomarkers to address the hierarchical organization of predictive models in the domain of executive functions.

As Diamond (2013) suggested, "without inhibitory control, we would be at the mercy of impulses, old habits of thought or action (conditioned responses), and/or stimuli in the environment that pull us this way or that". Computationally, within the predictive processing framework, these symptoms may assume a new meaning: not being able to evaluate the actual sensory stimulus against the internal model of a task, event, or in general, a context may result in the inability to recognize the salience of that stimulus for the current context. In the case of a stimulus that violates the internal model, being unable to recognize its salience can lead to a reduced or absent prediction error and, in turn, to the impossibility to update the internal model. From a behavioral point of view, this translates to the inability to initiate task-appropriate actions (Geng and Vossel, 2013) leading, for example, to perseverative errors and cognitive rigidity because of the firm adherence to the predictions or because of the persistence of premorbid predictions (Fotopoulou, 2014).

3.5 References

Barcelo, F., 2020. A predictive processing account of card sorting: Fast proactive and reactive frontoparietal cortical dynamics during inference and learning of perceptual categories. https://doi.org/10.31234/osf.io/zsw3t

Bardi, L., Six, P., Brass, M., 2017. Repetitive TMS of the temporo-parietal junction disrupts participant’s expectations in a spontaneous Theory of Mind task. Social Cognitive and Affective Neuroscience 12, 1775–1782. https://doi.org/10.1093/scan/nsx109

Blanke, O., Arzy, S., 2005. The Out-of-Body Experience: Disturbed Self-Processing at the Temporo-Parietal Junction. Neuroscientist 11, 16–24. https://doi.org/10.1177/1073858404270885

Buckner, R.L., Carroll, D.C., 2007. Self-projection and the brain. Trends in Cognitive Sciences 11, 49–57. https://doi.org/10.1016/j.tics.2006.11.004

Carter, R.M., Huettel, S.A., 2013. A nexus model of the temporal–parietal junction. Trends in Cognitive Sciences 17, 328–336. https://doi.org/10.1016/j.tics.2013.05.007

Corbetta, M., Patel, G., Shulman, G.L., 2008. The Reorienting System of the Human Brain: From Environment to Theory of Mind. Neuron 58, 306–324. https://doi.org/10.1016/j.neuron.2008.04.017

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3, 201–215. https://doi.org/10.1038/nrn755

Decety, J., Lamm, C., 2007. The Role of the Right Temporoparietal Junction in Social Interaction: How Low-Level Computational Processes Contribute to Meta-Cognition. Neuroscientist 13, 580–593. https://doi.org/10.1177/1073858407304654

Diamond, A., 2013. Executive Functions. Annu Rev Psychol 64, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dietz, M.J., Nielsen, J.F., Roepstorff, A., Garrido, M.I., 2021. Reduced effective connectivity between right parietal and inferior frontal cortex during audiospatial perception in neglect patients with a right-hemisphere lesion. Hearing Research, Stimulus-specific adaptation, MMN and predicting coding 399, 108052. https://doi.org/10.1016/j.heares.2020.108052

Doricchi, F., Pinto, M., Pellegrino, M., Marson, F., Aiello, M., Campana, S., Tomaiuolo, F., Lasaponara, S., 2021. Deficits of hierarchical predictive coding in left spatial neglect. Brain Communications 3, fcab111. https://doi.org/10.1093/braincomms/fcab111

Feldman, H., Friston, K., 2010. Attention, Uncertainty, and Free-Energy. Frontiers in Human Neuroscience 4, 215. https://doi.org/10.3389/fnhum.2010.00215

Fotopoulou, A., 2014. Time to get rid of the ‘Modular’ in neuropsychology: A unified theory of anosognosia as aberrant predictive coding. Journal of Neuropsychology 8, 1–19. https://doi.org/10.1111/jnp.12010

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Essen, D.C.V., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. PNAS 102, 9673–9678. https://doi.org/10.1073/pnas.0504136102

Garrido, M.I., Deouell, L.Y., 2021. Unilateral neglect within the predictive processing framework. Brain Communications 3, fcab193. https://doi.org/10.1093/braincomms/fcab193

Geng, J.J., Vossel, S., 2013. Re-evaluating the role of TPJ in attentional control: Contextual updating? Neuroscience & Biobehavioral Reviews 37, 2608–2620. https://doi.org/10.1016/j.neubiorev.2013.08.010

Godefroy, O., Martinaud, O., Narme, P., Joseph, P.-A., Mosca, C., Lhommée, E., Meulemans, T., Czernecki, V., Bertola, C., Labauge, P., Verny, M., Bellmann, A., Azouvi, P., Bindschaedler, C., Bretault, E., Boutoleau-Bretonniere, C., Robert, P., Lenoir, H., Krier, M., Roussel, M., GREFEX study group, 2018. Dysexecutive disorders and their diagnosis: A position paper. Cortex 109, 322–335. https://doi.org/10.1016/j.cortex.2018.09.026

Hohwy, J., 2013. The Predictive Mind. Oxford University Press.

Jakobs, O., Wang, L.E., Dafotakis, M., Grefkes, C., Zilles, K., Eickhoff, S.B., 2009. Effects of timing and movement uncertainty implicate the temporo-parietal junction in the prediction of forthcoming motor actions. NeuroImage 47, 667–677. https://doi.org/10.1016/j.neuroimage.2009.04.065

Kirsch, L.P., Mathys, C., Papadaki, C., Talelli, P., Friston, K., Moro, V., Fotopoulou, A., 2021. Updating beliefs beyond the here-and-now: the counter-factual self in anosognosia for hemiplegia. Brain Communications 3, fcab098. https://doi.org/10.1093/braincomms/fcab098

Klockgether, T., Dichgans, J., 1994. Visual control of arm movement in Parkinson's disease. Movement Disorders 9, 48–56. https://doi.org/10.1002/mds.870090108

Kocagoncu, E., Klimovich-Gray, A., Hughes, L.E., Rowe, J.B., 2020. Evidence and implications of abnormal predictive coding in dementia. arXiv:2006.06311 [q-bio].

Körding, K.P., Wolpert, D.M., 2004. Bayesian integration in sensorimotor learning. Nature 427, 244–247. https://doi.org/10.1038/nature02169

Macerollo, A., Palmer, C., Foltynie, T., Korlipara, P., Limousin, P., Edwards, M., Kilner, J.M., 2018. High-frequency peripheral vibration decreases completion time on a number of motor tasks. European Journal of Neuroscience 48, 1789–1802. https://doi.org/10.1111/ejn.14050

Moreira, H.S., Costa, A.S., Castro, S.L., Lima, C.F., Vicente, S.G., 2017. Assessing Executive Dysfunction in Neurodegenerative Disorders: A Critical Review of Brief

Neuropsychological Tools. Frontiers in Aging Neuroscience 9, 369. https://doi.org/10.3389/fnagi.2017.00369

Parr, T., Friston, K.J., 2018. The Computational Anatomy of Visual Neglect. Cerebral Cortex 28, 777–790. https://doi.org/10.1093/cercor/bhx316

Siman-Tov, T., Granot, R.Y., Shany, O., Singer, N., Hendler, T., Gordon, C.R., 2019. Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction. Neurosci Biobehav Rev 105, 262–275. https://doi.org/10.1016/j.neubiorev.2019.08.012

Wolpe, N., Zhang, J., Nombela, C., Ingram, J.N., Wolpert, D.M., Rowe, J.B., 2018. Sensory attenuation in Parkinson’s disease is related to disease severity and dopamine dose. Sci Rep 8, 15643. https://doi.org/10.1038/s41598-018-33678-3

Young, L., Dodell-Feder, D., Saxe, R., 2010. What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. Neuropsychologia 48, 2658–2664. https://doi.org/10.1016/j.neuropsychologia.2010.05.012

Study 2

1. Pre-stimulus standard alpha band power: supplementary figures

In this section we report additional images illustrating the main analysis on prestimulus standard alpha band power. Figures 1 and 2S depict the plots of the main interaction between time and prestimulus alpha power for the incongruent and congruent condition respectively, for significant electrodes.

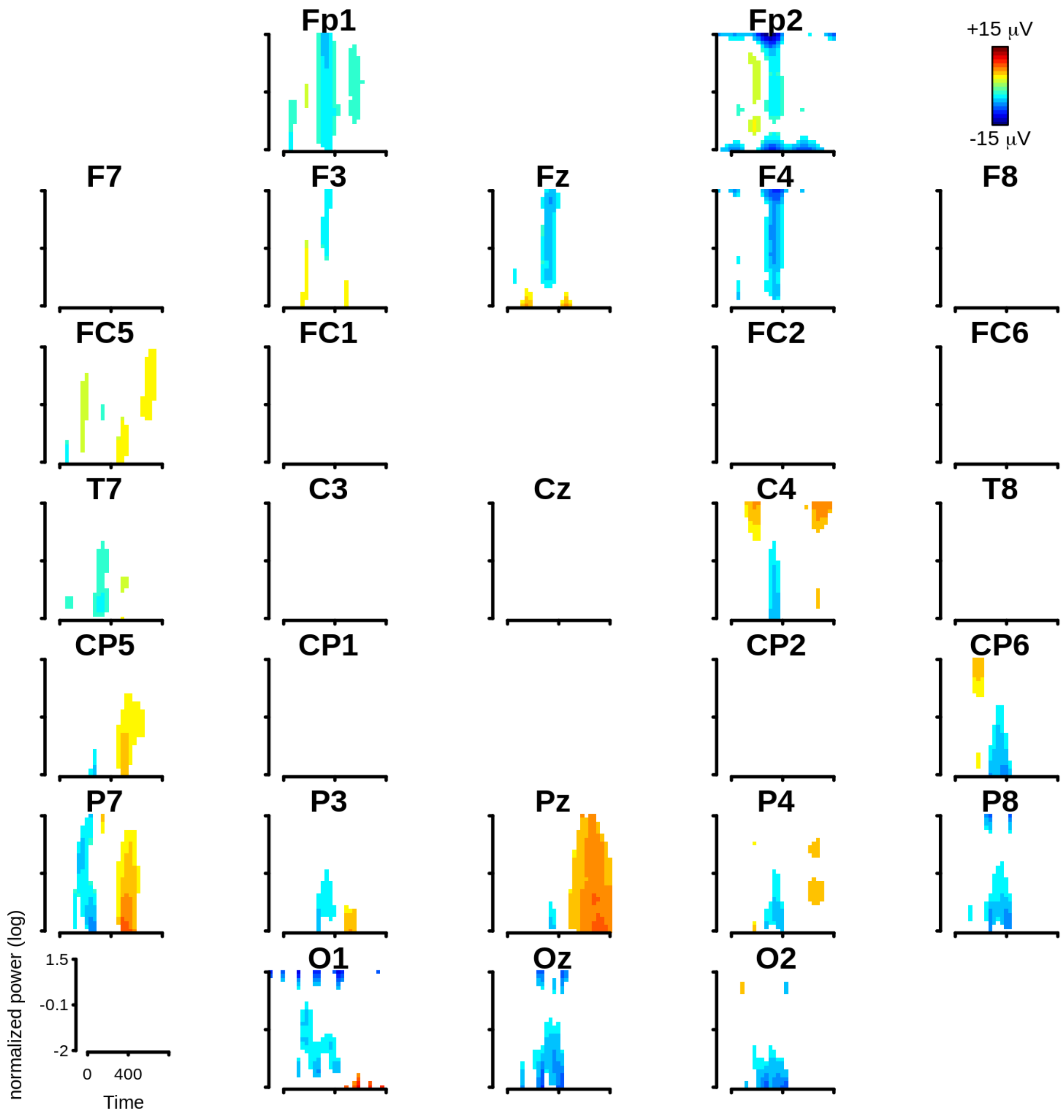


Figure 1S. 3-D plots of the main interaction between time and prestimulus alpha power for the incongruent condition, for significant electrodes. Time is represented on the x-axis, prestimulus power on the y-axis, and voltage amplitude is color coded. White areas indicate regions where the confidence intervals (95% CI) around the predicted surface included zero, while colored areas reflect amplitudes significantly predicted by prestimulus power. Shades of yellow and red indicate increasingly positive amplitudes, while shades of blue indicate increasingly negative amplitudes. In posterior electrodes, for lower levels of prestimulus power, the N400 was more negative while the P600 was more positive.

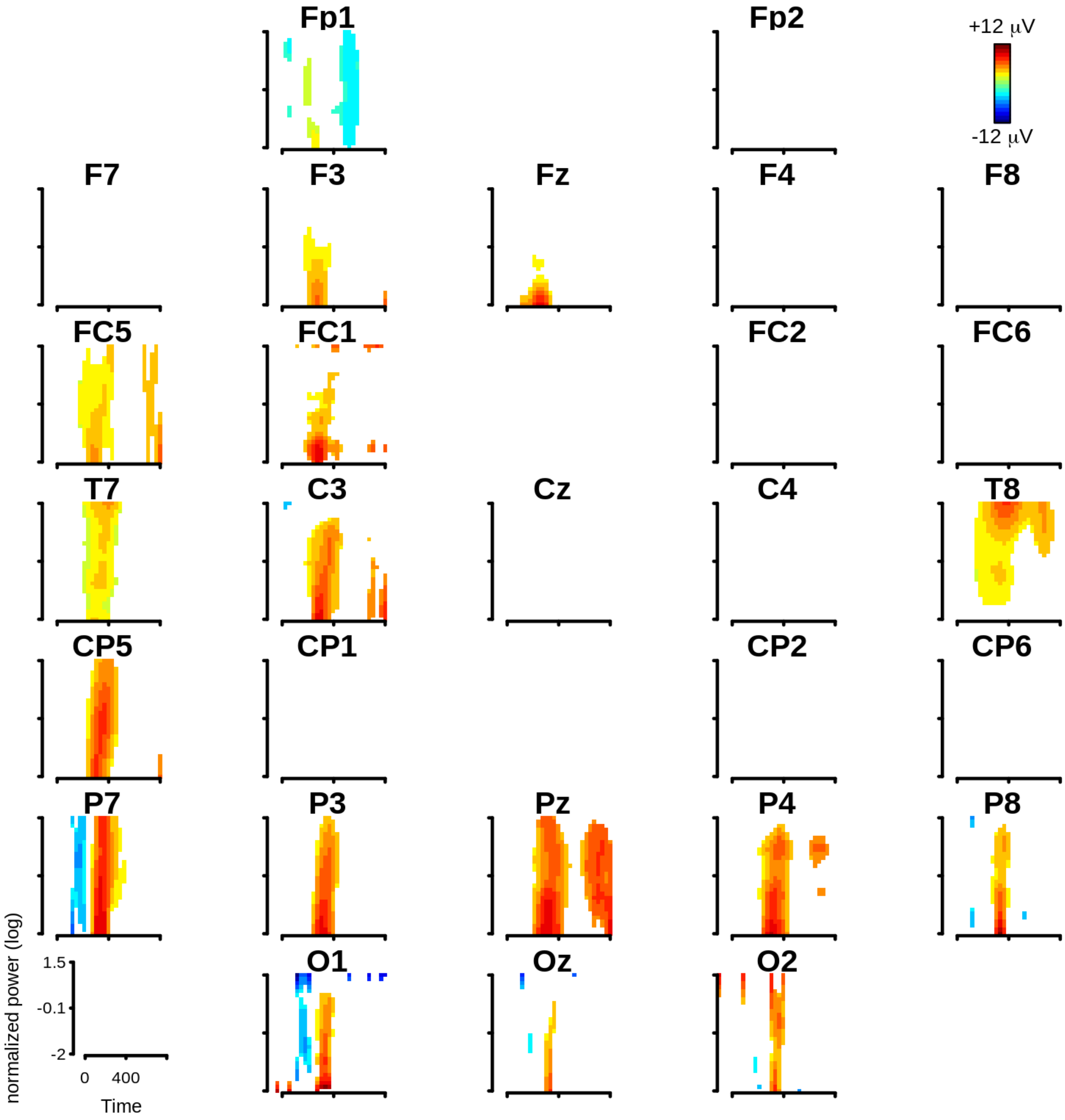


Figure 2S. 3-D plots of the main interaction between time and prestimulus alpha power for the congruent condition, for significant electrodes. Time is represented on the x-axis, prestimulus power on the y-axis, and voltage amplitude is color coded. White areas indicate regions where the confidence intervals (95% CI) around the predicted surface included zero, while colored areas reflect amplitudes significantly predicted by prestimulus power. Shades of yellow and red indicate increasingly positive amplitudes, while shades of blue indicate increasingly negative amplitudes. In central and posterior electrodes, voltage amplitude is more positive for lower levels of prestimulus alpha power.

Comparing these two figures we can see that lower prestimulus alpha power is predictive of enhanced voltage amplitude across conditions.

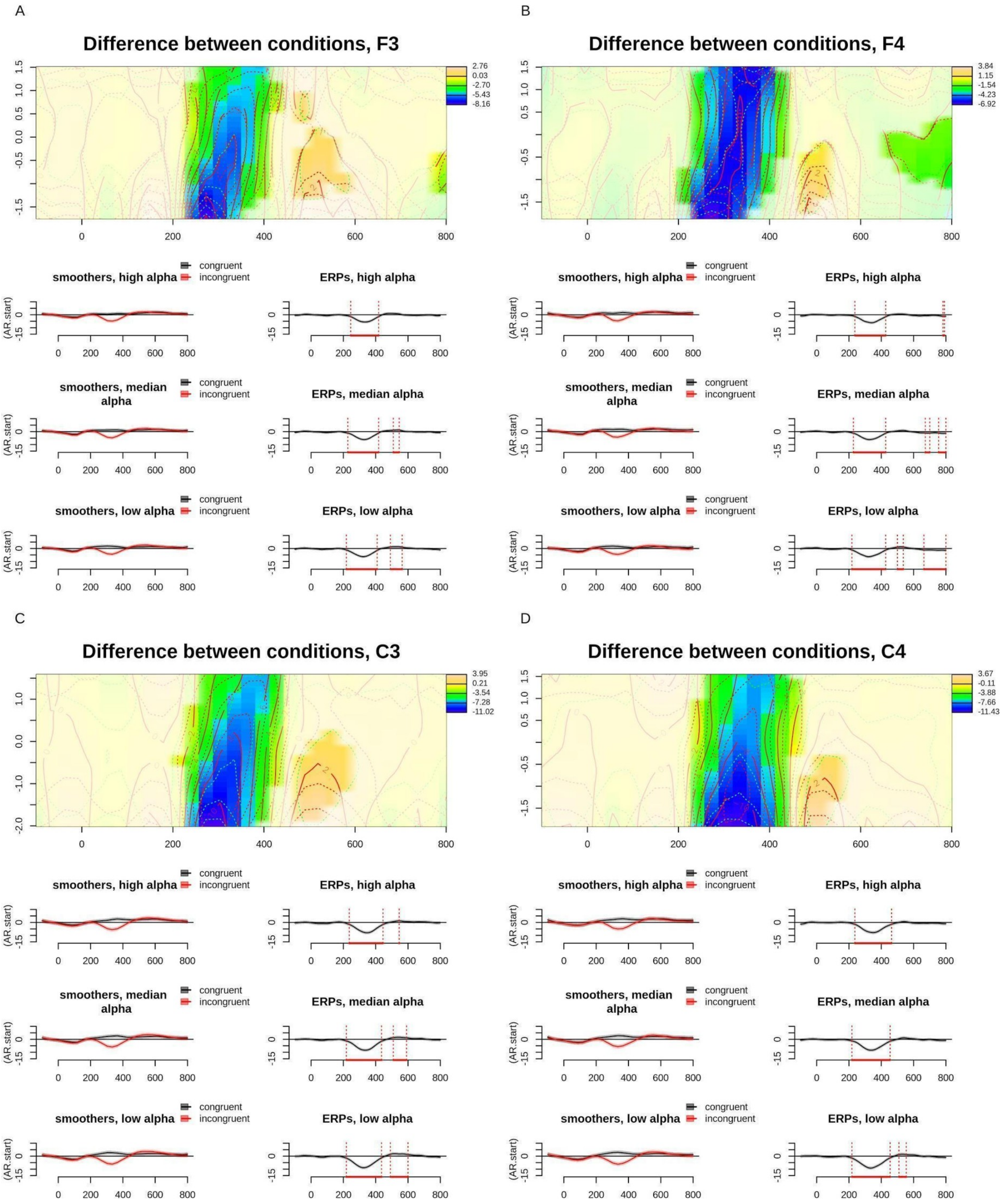


Figure 3S. Highlights for specific electrodes, for GAMMs on standard alpha band power. White areas indicate regions where the confidence intervals (95%) around the predicted surface included zero.r

2. Individual Alpha Peak Frequency results

In this section we report results for the analyses conducted on prestimulus individual alpha peak frequency (IAPF). Table 1S reports the significance of the interactions of interest for each electrode and condition. Figures 4 and 5S depict the effects of pre-stimulus IAPF on the ERPs.

Table 1S. Significance of interactions between time, power, and condition for the GAMM on IAPF power.

Electrodes	Interaction time x power x incongruent condition	Interaction: time x power x congruent condition
Fp1	0.01*	0.01*
Fp2	0.26	0.06
F7	0.39	0.39
F3	0.02*	0.01*
Fz	0.11	0.01*
F4	0.30	0.06
F8	0.10	0.89
FC5	< 0.001*	< 0.001*
FC1	0.16	0.01*
FC2	0.35	0.03*
FC6	0.35	0.47
T7	0.03*	0.01*
C3	0.21	< 0.001*
Cz	0.15	0.03*
C4	0.08	0.39
T8	0.40	< 0.001*
CP5	0.24	0.04*
CP1	0.55	0.01*
CP2	0.37	0.07
CP6	< 0.001*	0.09

P7	< 0.001*	< 0.001*
P3	0.02*	0.02*
Pz	0.04*	0.01*
P4	< 0.001*	0.06
P8	< 0.001*	< 0.001*
O1	0.01*	< 0.001*
Oz	0.01	< 0.001
O2	< 0.001	< 0.001

Significant results are marked with *.

As in Figure 4 in the main text, Figure 4S depicts, for each electrode, the tensor surface difference between the congruent and incongruent interactions of interest, i.e., the predicted amplitude difference between congruent and incongruent ERPs depending on prestimulus IAPF power. Colored blots represent significant differences, while white areas indicate regions where the confidence intervals (95% CI) around the predicted surface included zero. Shades of blue and green represent negative differences, while shades of red and yellow stand for positive differences.

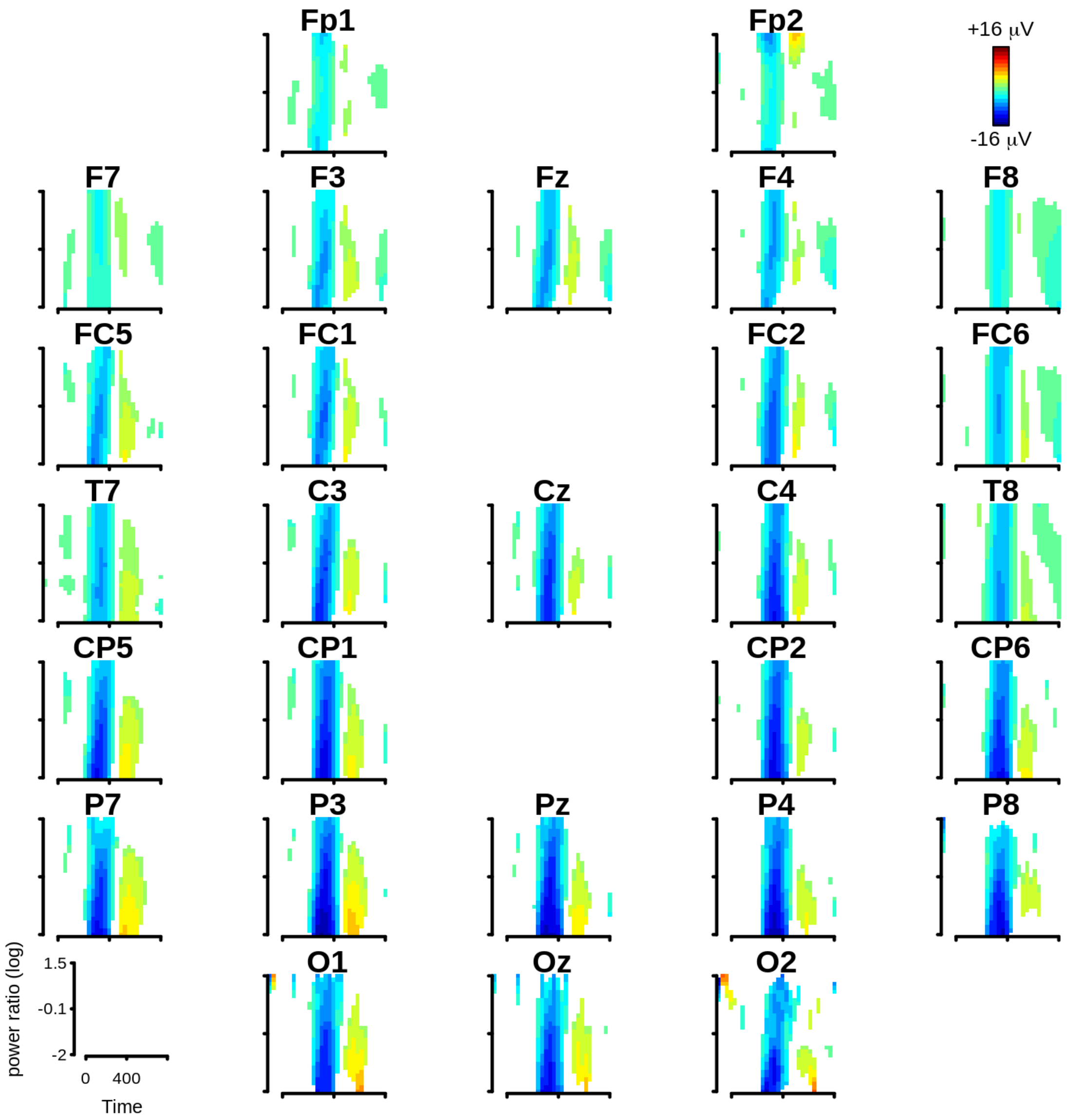


Figure 4S. Differences between incongruent and congruent tensor surfaces for the interaction between time, IAPF power, and condition.

Results are remarkably similar to those obtained with the standard alpha band with some differences, probably because of the exclusion of subjects without a clear peak in the 8-13 Hz range. In particular, similarly to the results obtained with the standard alpha frequency band, the predicted amplitude in the N400 and late posterior positivity/P600 time windows appeared to be significantly modulated by individualized prestimulus power, especially over central and posterior electrodes. To investigate whether the differences between the results were ascribable to the lower number of subjects included in the IAPF analyses, we conducted a follow-up exploratory analysis, running again the model on standard alpha frequency data, including only the 17 subjects included in the IAPF analyses. Results are shown in Figure 5S.

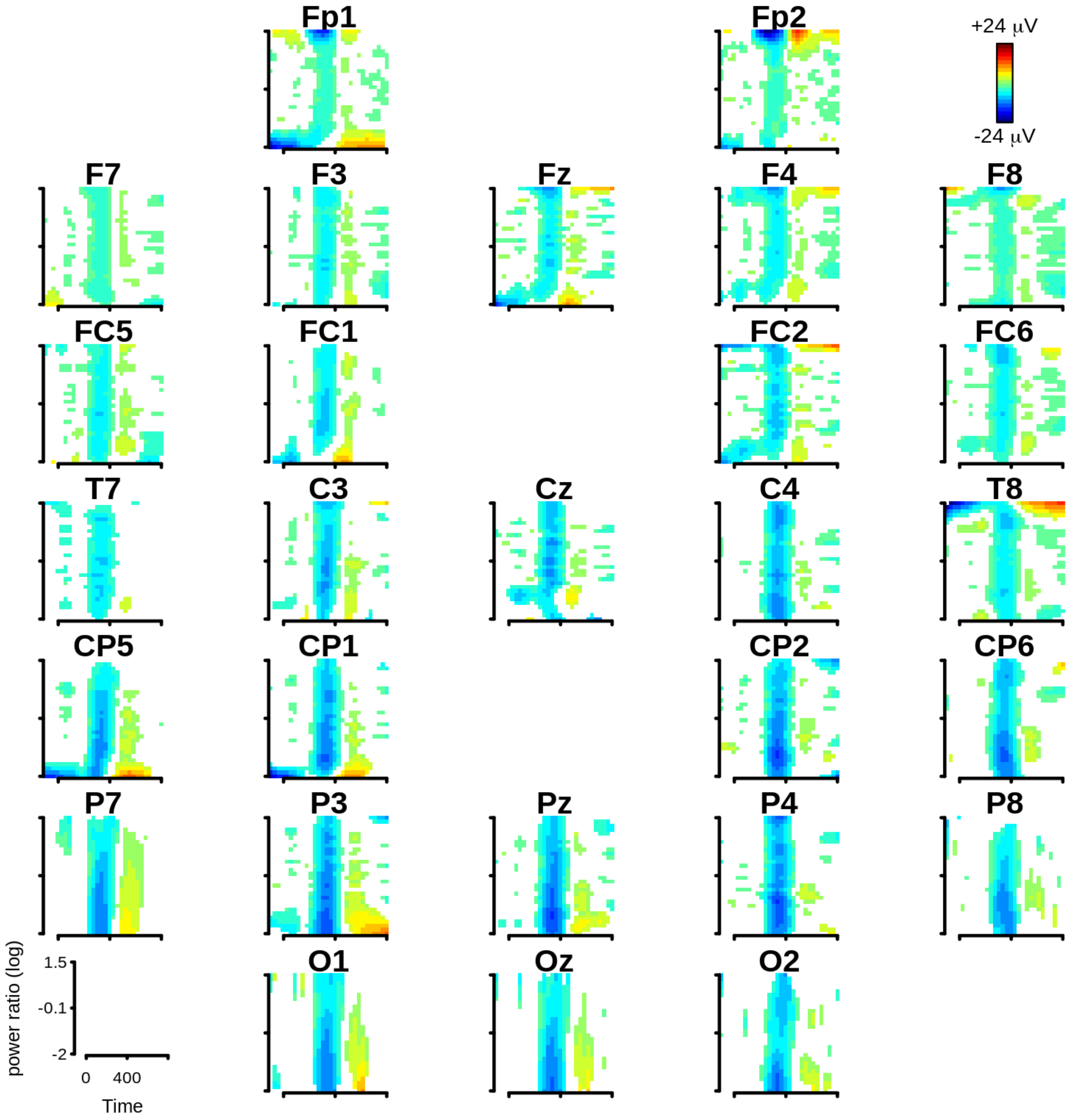
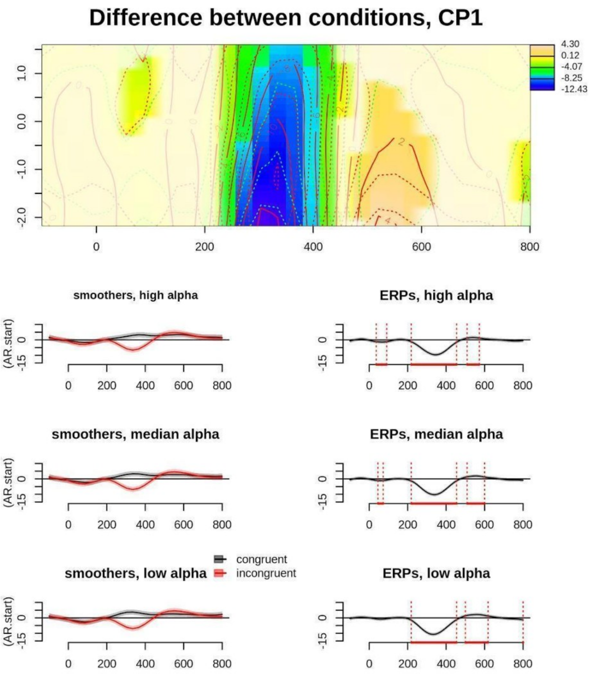


Figure 5S. Differences between incongruent and congruent tensor surfaces for the interaction between time, standard alpha band power, and condition (only for the 17 subjects included in the IAPF analyses).

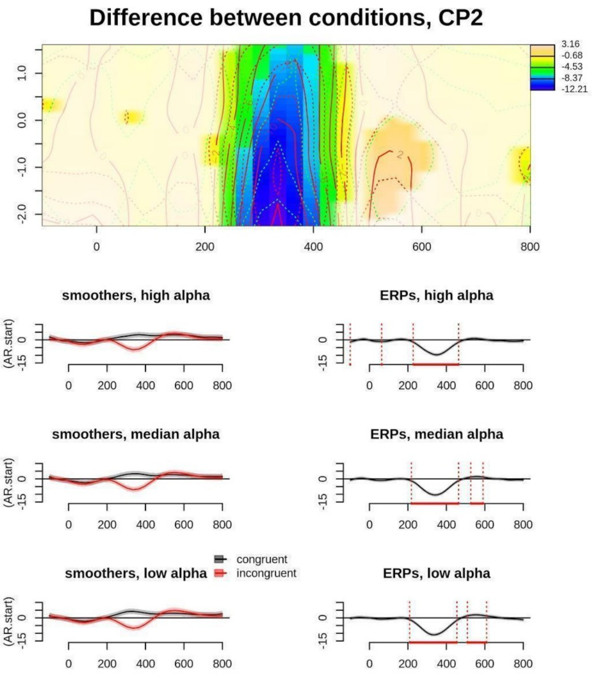
Comparing figures 4, 4S, and 5S we can see that figures 4S and 5S are very similar, albeit figure 4S represents IAPF results and figure 5S represents standard alpha band results, but only including the 17 subjects with a peak in the 8-13 Hz range. This seems to suggest that indeed that the differences between the results of standard alpha band power and IAPF power analyses are ascribable to the different number of subjects.

Significance values of the interaction between time, power, and congruent vs incongruent conditions for the results presented in this section, as well as highlights for specific electrodes, can be found in Table 2S and in Figures 6-9S.

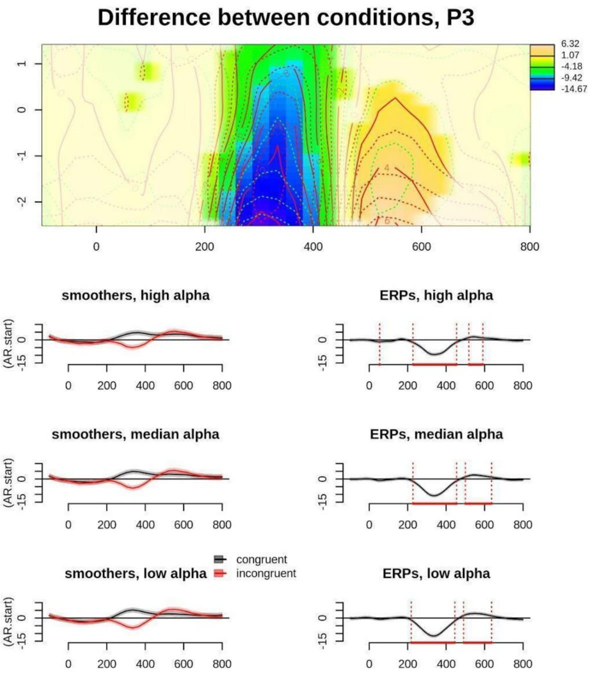
A


Difference between conditions, CP1

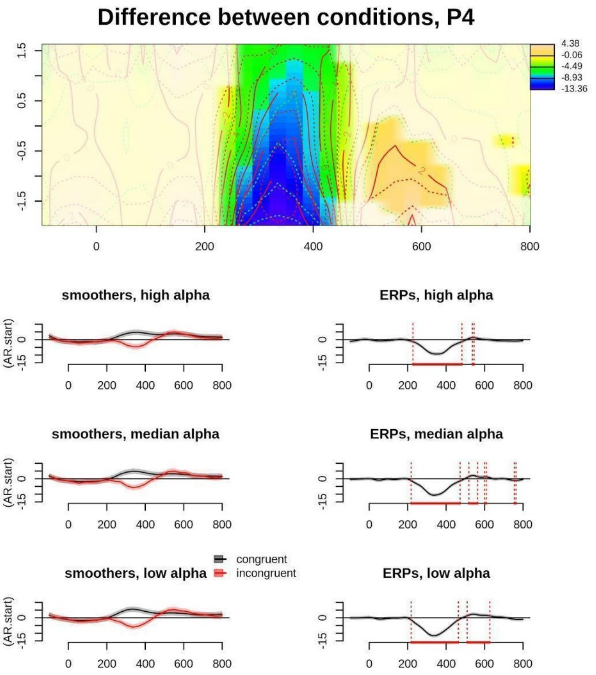
B


Difference between conditions, CP2

C


Difference between conditions, P3

D


Difference between conditions, P4

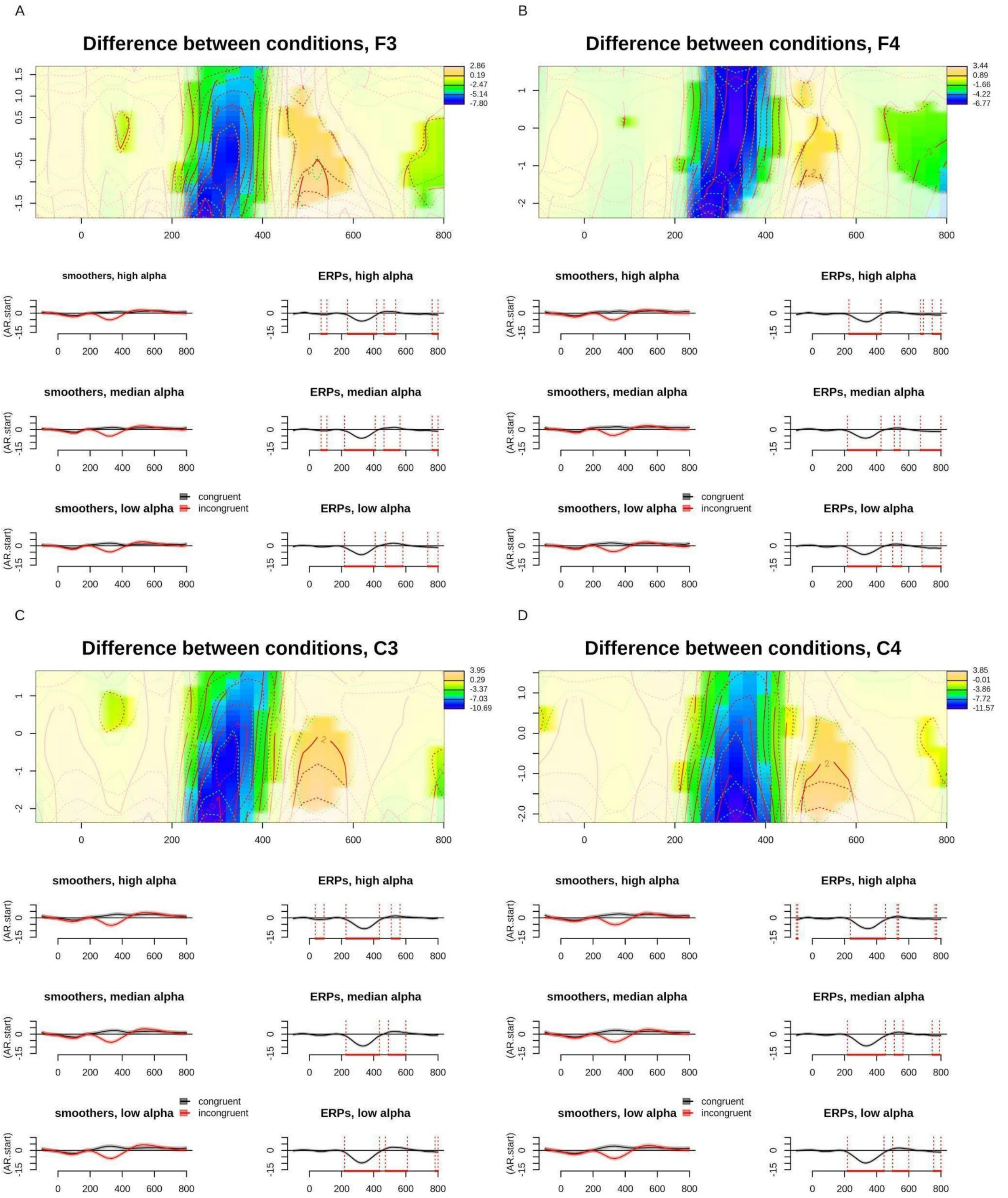


Figure 6-7S. Highlights for specific electrodes, for GAMMs on IAPF power. White areas indicate regions where the confidence intervals (95%) around the predicted surface included zero.

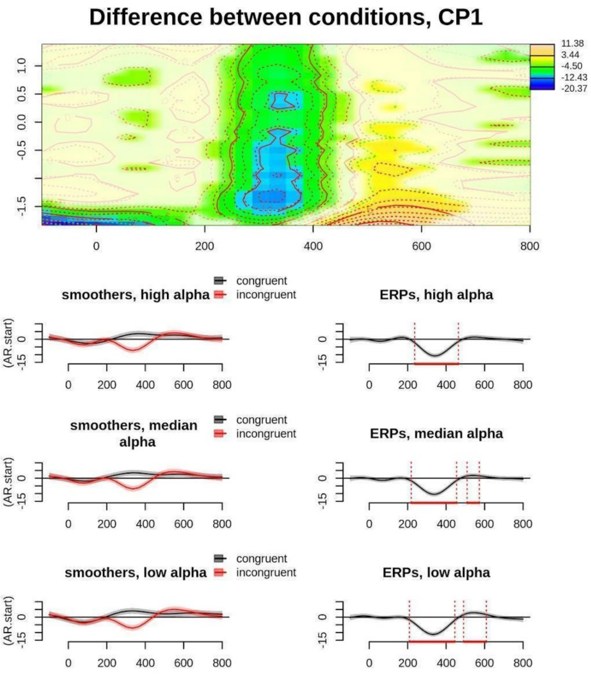
Table 2S. Significance of interactions between time, power, and condition for the GAMM on the 17 subjects included in the IAPF analyses.

Electrodes	Interaction time x power x	Interaction time x power x

	incongruent condition	congruent condition
Fp1	< 0.001*	< 0.001*
Fp2	< 0.001*	< 0.001*
F7	< 0.001*	< 0.001*
F3	< 0.001*	< 0.001*
Fz	< 0.001*	< 0.001*
F4	< 0.001*	< 0.001*
F8	< 0.001*	< 0.001
FC5	< 0.001*	< 0.001*
FC1	< 0.001*	< 0.001*
FC2	< 0.001*	< 0.001*
FC6	< 0.001*	0.41
T7	< 0.001*	< 0.001*
C3	< 0.001*	0.02*
Cz	< 0.001*	< 0.001*
C4	< 0.001*	0.15
T8	< 0.001*	< 0.001*
CP5	< 0.001*	< 0.001*
CP1	< 0.001*	< 0.001*
CP2	< 0.001*	< 0.001*
CP6	< 0.001*	< 0.001*
P7	< 0.001*	0.24
P3	< 0.001*	< 0.001*
Pz	< 0.001*	< 0.001*
P4	< 0.001*	< 0.001*
P8	< 0.001*	< 0.001*
O1	< 0.001*	< 0.001*
Oz	< 0.001*	< 0.001*
O2	< 0.001*	< 0.001*

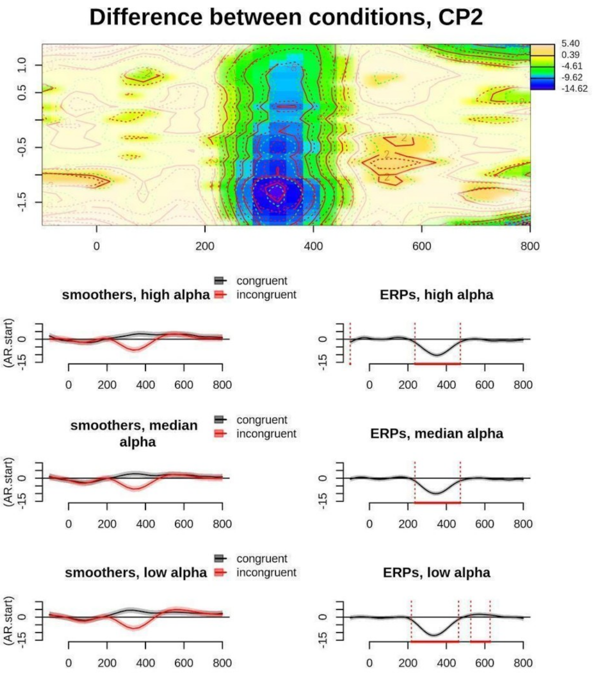
Significant results are marked with *.

A



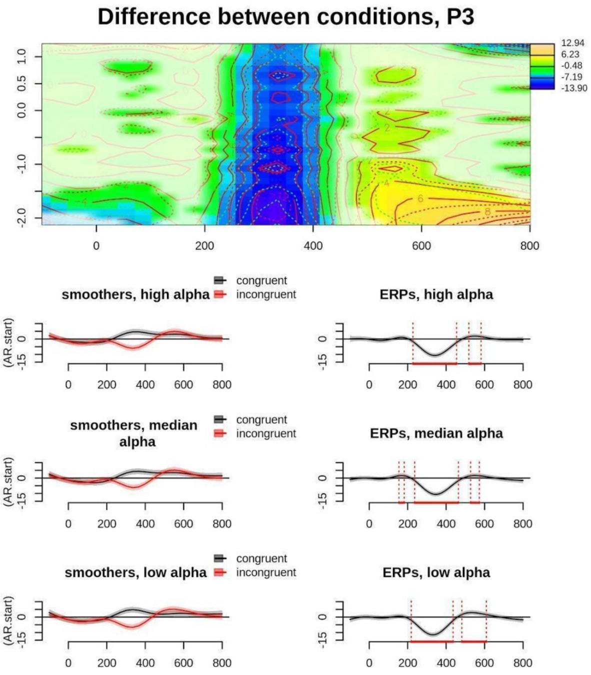
Difference between conditions, CP1
11.38
3.44
-4.50
-12.43
-20.37
smoothers, high alpha
ERPs, high alpha
smoothers, median alpha
ERPs, median alpha
smoothers, low alpha
ERPs, low alpha
congruent
incongruent
(AR.start)

B



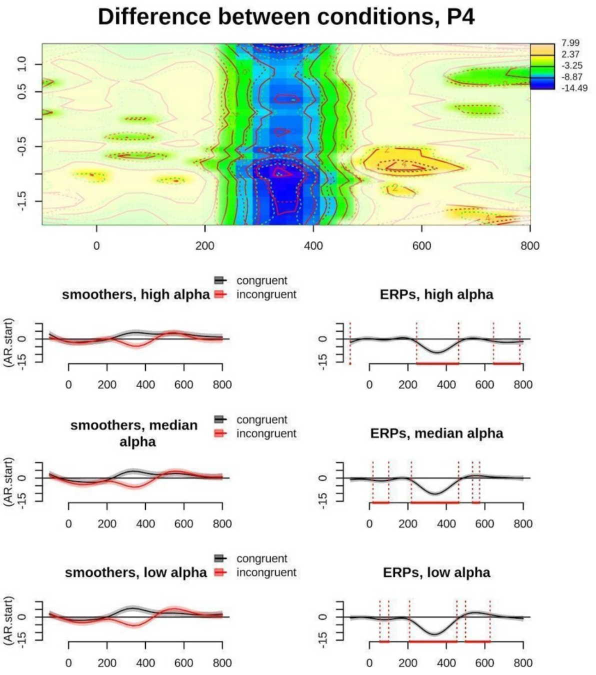
Difference between conditions, CP2
5.40
0.39
-4.61
-9.62
-14.62
smoothers, high alpha
ERPs, high alpha
smoothers, median alpha
ERPs, median alpha
smoothers, low alpha
ERPs, low alpha
congruent
incongruent
(AR.start)

C

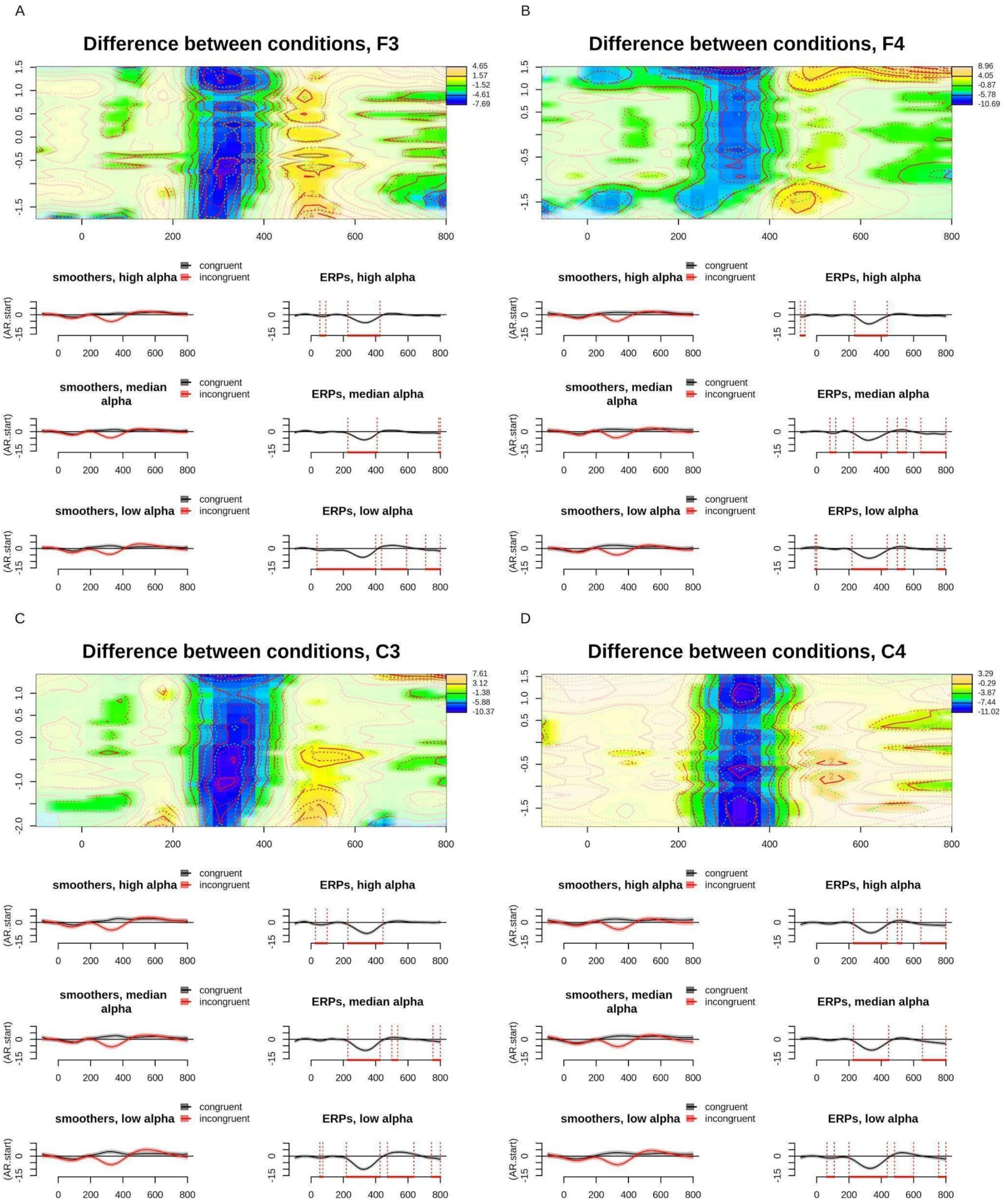


Difference between conditions, P3
12.94
6.23
-0.48
-7.19
-13.90
smoothers, high alpha
ERPs, high alpha
smoothers, median alpha
ERPs, median alpha
smoothers, low alpha
ERPs, low alpha
congruent
incongruent
(AR.start)

D



Difference between conditions, P4
7.99
2.37
-3.25
-8.87
-14.49
smoothers, high alpha
ERPs, high alpha
smoothers, median alpha
ERPs, median alpha
smoothers, low alpha
ERPs, low alpha
congruent
incongruent
(AR.start)



Figures 8-9S. Highlights for specific electrodes, for GAMMs on the standard alpha band, only including the 17 subjects included in the IAPF analyses. White areas indicate regions where the confidence intervals (95%) around the predicted surface included zero.

Standard alpha band power has been studied more extensively and linked to a variety of cognitive processes (see Section 1), but recently introduced methodologies allow an easier estimation of multiple properties of alpha oscillations. For example, higher IAPF (i.e., faster alpha band oscillations) seems to be related to the sampling resolution of visual information (Samaha and Postle, 2015), probably enabling faster processing speed and therefore optimizing the efficiency of basic cognitive functions such as attention. In this vein, a positive correlation between high IAPF and general intelligence was found (Grandy et al., 2013), while in patients with schizophrenia, lower IAPF has been associated with poorer discrimination of visual targets in an attention task, as well as impaired global cognition when compared to healthy controls (Ramsay et al., 2021). These results substantiate the hypothesis that higher IAPF optimizes the efficacy of basic cognitive functions, with cascading positive effects also on higher cognitive abilities. However, there are also reports contrasting this view: Howard et al., 2017 found an advantage for individuals with a lower IAPF in a spatial attention task, and Ociepka et al., 2022 found that general intelligence was unrelated to the IAPF, but correlated positively with processing speed, leading to the conclusion that "brains with higher IAFs do run faster, but it does not make them smarter".

In the language domain, it seems that overall, a lower IAPF is associated with more efficient linguistic processing: for example, lower-IAPF individuals show a better performance than their higher-IAPF counterparts when learning a modified miniature language (Nalaye et al., 2022). As regards the ERP correlates of language processing instead, low-IAPF individuals showed a sustained positivity when processing syntactically or semantically ambiguous sentences, while high-IAPF individuals did not (Bornkessel et al., 2004). In another study, the authors examined whether the speed of predictive model adaptation in response to intra-experimental probabilistic information, reflected in the N400 amplitudes, were different between high and low-IAPF participants. Interestingly, participants with lower IAPF showed a faster model adaptation (Bornkessel-Schlesewsky et al., 2022).

This body of evidence is in line with our findings on IAPF power, but it suggests also that other properties of alpha oscillations, such as IAPF speed, might be predictive of post-stimulus cognitive processes and electrophysiological responses. This field therefore deserves further investigation.

3. Additional analyses

To explore whether a more traditional analysis would have captured the modulations associated with prestimulus alpha we performed some additional analyses using ANOVAs instead of GAMM. In particular, we performed two separate ANOVAs, one on the N400 and one on the P600 time window. To investigate topographic effects, we followed the same method as in the original paper (Arcara et al., 2019): we selected 12 electrodes grouped in 4 Regions of interest (ROI): a left anterior (F3, FC5, FC1) a right anterior (F4, FC6, FC2), a left posterior (CP1, CP5, P3) and a right posterior (CP2, CP6, P4). ERP amplitudes for each ROI were calculated as the mean amplitude of the electrodes in the two time windows of interest, based on the literature: the first from 300 to 500 post-stimulus (Kuperberg et al., 2020), to capture the N400 effect; and the second from 500 to 700 msec post-stimulus, to capture the P600 (Mancini et al., 2019).

To investigate the effects of prestimulus alpha on the ERP amplitudes, we averaged every single-trial raw prestimulus alpha value over the electrodes of each ROI. We then calculated two individualized cut-off values for prestimulus alpha power, corresponding to the first and third quartile of each participant’s set of prestimulus power extracted from every single trial (Arcara et al., 2021). The trials were then labelled as “high power” and “low power” according to whether their power value fell above the third or below the first quartile. Trials falling in between the two quartiles were discarded from the analysis. In this way, each ERP amplitude value was associated with a label indicating whether that ERP was preceded by high or low prestimulus power. The repeated ANOVAs included three within variables with a 2 × 2 × 4 design: Congruence with two levels (congruent and incongruent); Prestimulus

Alpha Power Level with two levels (high and low alpha) and Position with four levels (left anterior, right anterior, left posterior and right posterior ROIs). Since more than two levels were involved in the Position variable, a preliminary Mauchly test for sphericity was performed. If sphericity assumption was not met, Greenhouse-Geisser correction was applied. Effect size for ANOVA effects was calculated as global eta squared (η^2_G), which is a more accurate estimate of effect size than traditional η^2_p in the case of repeated measure design (Bakeman, 2005). Post-hoc contrasts were performed by means of paired t-tests, corrected for multiple comparisons with Bonferroni correction method.

3.1. Additional analyses results

Contrast on congruence, N400 time window (300-500). In this analysis we found a significant interaction of congruence × position [$F(3,60) = 25.85$, $p < .05$, $\eta^2_G = .56$]. Post-hoc contrasts related to this interaction showed that over all ROIs, incongruent trials elicited more negative ERPs than congruent ones in the N400 time window (corrected ps < .05; see representative electrodes for the ROIs in figure A1). All results for these contrasts are shown in Tables from 3A to 10A.

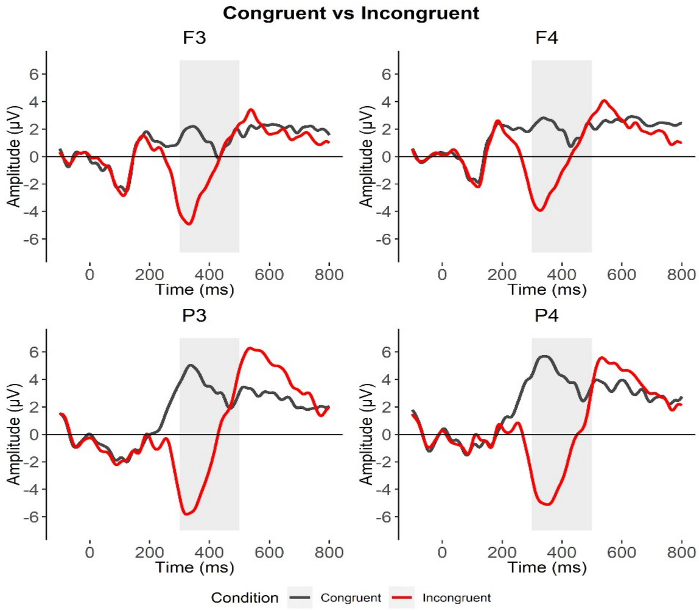


Figure 10S. Results of repeated measures ANOVAs in the N400 time window (300-500 ms).

Contrast on congruence, P600 time window (500-700). In this analysis we found a significant interaction of congruence × position [F(3,60) = 8.17, $p < .05$, $\eta^2_G = .29$]. Post-hoc contrasts related to this interaction showed that in the left posterior ROIs (electrode P3 in Figure A2), incongruent trials elicited a more positive ERP in the P600 time window than congruent trials. All results for these contrasts are shown in Tables from A.2.1. to A.2.4.

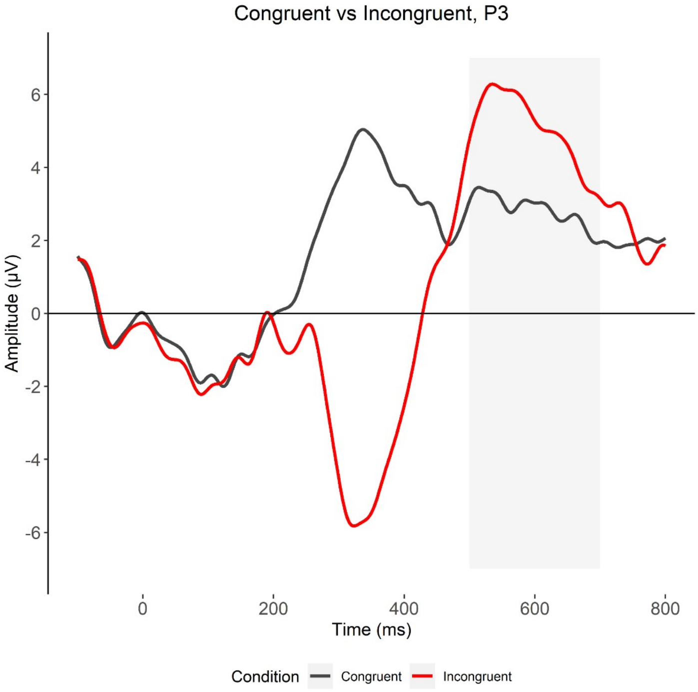


Figure 11S. Results of repeated measures ANOVAs in the P600 time window (500-700 ms).

The absence of significant interactions between congruence and prestimulus alpha level, or between congruence, prestimulus power and position for the P600 time-window could seem somewhat surprising, especially if compared to the GAMM results in this time-window. However, we think that the reason for the lack of significant interactions between congruence and prestimulus power (and position) lies in that ANOVA testing is not suitable to detect such fine-grained effects, since this statistical technique requires the prior averaging and factorization of data before entering it in the model (see Section 1.3). This leads to the loss of important variability within the data, that negatively influences the statistical outcomes.

Table 3S. Results of ANOVA of contrast on Congruence, N400 time window (300-500 msec). This table reports the full ANOVA results. The first column reports the effect name, the second and the third the degrees of freedom of numerator and denominator. The fourth column reports the F value. The fifth column reports the p-value, and the sixth column reports an asterisk if the p.value of the effect is below 0.05. The final column reports the ηG2 (global eta squared), a measure of effect size (Bakeman, 2005).

Contrast on Congruency, N400 time window (300-500 msec) ANOVA detailed results						
ANOVA terms	DFn	DFd	F	p-value	p<0.05	ges
Congruence	1.00	20.00	177.41	< 0.001	*	0.90
Prestimulus_alpha	1.00	20.00	1.47	0.24		0.07
Position	3.00	60.00	3.83	0.01	*	0.16
Congruence x Prestimulus_alpha	1.00	20.00	< 0.001	0.98		< 0.001
Congruence x Position	3.00	60.00	25.85	< 0.001	*	0.56
Prestimulus_alpha x Position	3.00	60.00	2.80	0.05	*	0.12
Congruence x Prestimulus_alpha:Position	3.00	60.00	2.63	0.06		0.12

Table 4S. Results of Mauchly's test associated with ANOVA for contrast on Congruency, N400 time window (300-500 ms). This table reports the results of Mauchly's test associated with ANOVA. The first column reports the effect name, the second the W value, the third the p-value, the fourth an asterisk if the p-value associated with the Mauchly's test was below 0.05 (i.e., sphericity assumption violated).

<table>
<tr><td colspan="4">Contrast on Congruency, N400 time window (300-500 msec)
Mauchly’s test</td></tr>
<tr><td>Mauchly’s test for Sphericity</td><td>W</td><td>p-value</td><td>p<0.05</td></tr>
<tr><td>position</td><td>0.40</td><td>< 0.001</td><td>*</td></tr>
<tr><td>Cond:position</td><td>0.69</td><td>0.22</td><td></td></tr>
<tr><td>trial_type:position</td><td>0.29</td><td>0.01</td><td>*</td></tr>
<tr><td>Cond:trial_type:position</td><td>0.71</td><td>0.27</td><td></td></tr>
</table>

Table 5S. Results of Sphericity corrections associated with ANOVA for contrast on Congruence, N400 time window (300- 500 ms). This table reports the sphericity corrections associated with the contrast. The first column reports the effect names, the second column reports the Greenhouse-Geisser (GG) epsilon. The third reports the p-value with GG correction, the fourth column reports if the p-value with GG correction was below 0.05. The fifth column reports the value of the Huynh-Feldt (HF) epsilon. The sixth the p-values with HF correction, and the seventh reports an asterisk if the p-value with HF correction is below 0.05.

<table>
<tr><td colspan="7">Contrast on Congruency, N400 time window (300-500 msec)
Sphericity corrections of p-values</td></tr>
<tr><td>Sphericity correction terms</td><td>GGe</td><td>p[GG]</td><td>p[GG] < .05</td><td>HFe</td><td>p[HF]</td><td>p[HF]<.05</td></tr>
<tr><td>position</td><td>0.65</td><td>0.03</td><td>*</td><td>0.71</td><td>0.03</td><td>*</td></tr>
<tr><td>Cond:position</td><td>0.79</td><td>0.01</td><td>*</td><td>0.91</td><td>0.01</td><td>*</td></tr>
<tr><td>trial_type:position</td><td>0.64</td><td>0.08</td><td></td><td>0.71</td><td>0.07</td><td></td></tr>
</table>

Cond:trial_type:position	0.82	0.07		0.95	0.06	

Table 6S. Post-hocs for the term Congruence x Position (contrast on Congruence, N400 time window, 300-500 ms). The table reports all the post-hocs. The first column reports the pairwise contrast. The second column reports the t value. The third column reports the p-value, corrected with the Bonferroni method. The fourth column reports the mean of the first term. The fifth column reports the mean of the second term. The sixth column marks with * significant contrasts.

post-hocs (contrast on Congruence, N400 time window) Congruence x Position					
comparison	t-value (df=24)	p-value	Mean 1	Mean 2	p<.05
CON_left_anterior vs CON_left_posterior	-5.37	< 0.001	1.65	3.11	*
CON_left_anterior vs CON_right_anterior	-3.20	0.13	1.65	2.39	
CON_left_anterior vs CON_right_posterior	-6.58	< 0.001	1.65	4.02	*
CON_left_anterior vs INC_left_anterior	7.12	< 0.001	1.65	-1.25	*
CON_left_anterior vs INC_left_posterior	8.22	< 0.001	1.65	-1.86	*
CON_left_anterior vs INC_right_anterior	6.30	< 0.001	1.65	-0.92	*
CON_left_anterior vs INC_right_posterior	7.07	< 0.001	1.65	-1.86	*
CON_left_posterior vs CON_right_anterior	2.12	1.00	3.11	2.39	
CON_left_posterior vs CON_right_posterior	-4.52	0.01	3.11	4.02	*

CON_left_posterior vs INC_left_anterior	11.05	< 0.001	3.11	-1.25	*
CON_left_posterior vs INC_left_posterior	13.83	< 0.001	3.11	-1.86	*
CON_left_posterior vs INC_right_anterior	8.99	< 0.001	3.11	-0.92	*
CON_left_posterior vs INC_right_posterior	11.95	< 0.001	3.11	-1.86	*
CON_right_anterior vs CON_right_posterior	-4.67	< 0.001	2.39	4.02	*
CON_right_anterior vs INC_left_anterior	10.06	< 0.001	2.39	-1.25	*
CON_right_anterior vs INC_left_posterior	9.57	< 0.001	2.39	-1.86	*
CON_right_anterior vs INC_right_anterior	8.53	< 0.001	2.39	-0.92	*
CON_right_anterior vs INC_right_posterior	8.33	< 0.001	2.39	-1.86	*
CON_right_posterior vs INC_left_anterior	12.81	< 0.001	4.02	-1.25	*
CON_right_posterior vs INC_left_posterior	14.15	< 0.001	4.02	-1.86	*
CON_right_posterior vs INC_right_anterior	10.24	< 0.001	4.02	-0.92	*
CON_right_posterior vs INC_right_posterior	13.37	< 0.001	4.02	-1.86	*
INC_left_anterior vs INC_left_posterior	1.75	1.00	-1.25	-1.86	
INC_left_anterior vs INC_right_anterior	-1.22	1.00	-1.25	-0.92	
INC_left_anterior vs INC_right_posterior	1.43	1.00	-1.25	-1.86	
INC_left_posterior vs INC_right_anterior	-2.72	0.37	-1.86	-0.92	

INC_left_posterior vs INC_right_posterior	< 0.001	1.00	-1.86	-1.86	
INC_right_anterior vs INC_right_posterior	2.54	0.54	-0.92	-1.86	

Table 7S. Results of ANOVA of contrast on Congruence, P600 time window (500-700 msec). This table reports the full ANOVA results. The first column reports the effect name, the second and the third the degrees of freedom of numerator and denominator. The fourth column reports the F value. The fifth column reports the p-value, and the sixth column reports an asterisk if the p.value of the effect is below 0.05. The final column reports the ηG2 (global eta squared), a measure of effect size (Bakeman, 2005).

Contrast on Congruency, P600 time window (500-700 msec) ANOVA detailed results						
ANOVA terms	DFn	DFd	F	p-value	p<0.05	ges
Congruence	1.00	20.00	6.98	0.02	*	0.26
Prestimulus_alpha	1.00	20.00	2.25	0.15		0.10
Position	3.00	60.00	12.19	< 0.001	*	0.38
Congruence:Prestimulus_alpha	1.00	20.00	2.74	0.11		0.12
Congruence:Position	3.00	60.00	8.17	< 0.001	*	0.29
Prestimulus_alpha:Position	3.00	60.00	0.18	0.91		0.01
Congruence:Prestimulus_alpha:Position	3.00	60.00	0.16	0.92		0.01

Table 8S. Results of Mauchly's test associated with ANOVA for contrast on Congruency, P600 time window (500-700 msec). This table reports the results of Mauchly's test associated with ANOVA. The first column reports the effect name, the second the W value, the third the p-value, the fourth an asterisk if the p-value associated with the Mauchly's test was below 0.05 (i.e., sphericity assumption violated).

Contrast on Congruency, N400 time window (300-500 msec) Mauchly's test			
Mauchly's test for Sphericity	W	p-value	p<0.05
Position	0.42	0.01	*
Congruence:Position	0.54	0.04	*
Prestimulus_alpha:Position	0.29	< 0.001	*
Congruence:Prestimulus_alpha:Position	0.44	0.01	*

Table 9S. Results of Sphericity corrections associated with ANOVA for contrast on Congruence, P600 time window (500-700 msec). This table reports the sphericity corrections associated with the contrast. The first column reports the effect names, the second column reports the Greenhouse-Geisser (GG) epsilon. The third reports the p-value with GG correction, the fourth column reports if the p-value with GG correction was below 0.05. The fifth column reports the value of the Huynh-Feldt (HF) epsilon. The sixth the p-values with HF correction, and the seventh reports an asterisk if the p-value with HF correction is below 0.05.

Contrast on Congruency, P600 time window (500-700 msec) Sphericity corrections of p-values						
Sphericity correction terms	GGe	p[GG]	p[GG]	HFe	p[HF]	p[HF]<.05

			< .05			
Position	0.65	< 0.001	*	0.71	< 0.001	*
Congruence:Position	0.77	< 0.001	*	0.88	< 0.001	*
Prestimulus_alpha:Position	0.57	0.80		0.62	0.82	
Congruence:Prestimulus_alpha:Position	0.65	0.85		0.72	0.87	

Table 10S. Post-hocs for the term Congruence x Position (contrast on Congruence, P600 time window, 500-700 msec). The table reports all the post-hocs. The first column reports the pairwise contrast. The second column reports the t value. The third column reports the p-value, corrected with the Bonferroni method. The fourth column reports the mean ofthe first term. The fifth column reports the mean of the second term. The sixth column marks with * significant contrasts.

post-hocs (contrast on Congruence, P600 time window) Congruence x Position					
comparison	t-value (df=24)	p-value	Mean 1	Mean 2	p<.05
CON_left_anterior vs CON_left_posterior	-1.81	1.00	2.06	2.63	
CON_left_anterior vs CON_right_anterior	-2.55	0.53	2.06	2.64	
CON_left_anterior vs CON_right_posterior	-3.11	0.15	2.06	3.09	
CON_left_anterior vs INC_left_anterior	-1.04	1.00	2.06	2.50	
CON_left_anterior vs INC_left_posterior	-4.68	< 0.001	2.06	4.49	*

CON_left_anterior vs INC_right_anterior	-2.17	1.00	2.06	2.71	
CON_left_anterior vs INC_right_posterior	-4.28	0.01	2.06	4.00	*
CON_left_posterior vs CON_right_anterior	-0.01	1.00	2.63	2.64	
CON_left_posterior vs CON_right_posterior	-2.46	0.65	2.63	3.09	
CON_left_posterior vs INC_left_anterior	0.31	1.00	2.63	2.50	
CON_left_posterior vs INC_left_posterior	-4.71	< 0.001	2.63	4.49	*
CON_left_posterior vs INC_right_anterior	-0.19	1.00	2.63	2.71	
CON_left_posterior vs INC_right_posterior	-3.43	0.07	2.63	4.00	
CON_right_anterior vs CON_right_posterior	-1.23	1.00	2.64	3.09	
CON_right_anterior vs INC_left_anterior	0.29	1.00	2.64	2.50	
CON_right_anterior vs INC_left_posterior	-3.62	0.05	2.64	4.49	*
CON_right_anterior vs INC_right_anterior	-0.20	1.00	2.64	2.71	
CON_right_anterior vs INC_right_posterior	-2.82	0.30	2.64	4.00	
CON_right_posterior vs INC_left_anterior	1.58	1.00	3.09	2.50	
CON_right_posterior vs INC_left_posterior	-3.54	0.06	3.09	4.49	
CON_right_posterior vs INC_right_anterior	1.03	1.00	3.09	2.71	
CON_right_posterior vs INC_right_posterior	-2.36	0.79	3.09	4.00	

INC_left_anterior vs INC_left_posterior	-6.36	< 0.001	2.50	4.49	*
INC_left_anterior vs INC_right_anterior	-0.84	1.00	2.50	2.71	
INC_left_anterior vs INC_right_posterior	-4.18	0.01	2.50	4.00	*
INC_left_posterior vs INC_right_anterior	4.38	0.01	4.49	2.71	*
INC_left_posterior vs INC_right_posterior	1.33	1.00	4.49	4.00	
INC_right_anterior vs INC_right_posterior	-4.32	0.01	2.71	4.00	*

4. Stimulus properties and spontaneous fluctuations

In the context of the present study, we hypothesized that the observed pre-stimulus alpha power is related to spontaneous fluctuations in a physiological state, not related to any external manipulations. It is important to note that it is very difficult to demonstrate that these are actual spontaneous fluctuations and are not instead related to some cue-target pair's properties. In short, since all stimuli were not equal, it could be the case that some specific stimuli were associated to lower prestimulus alpha and in turn to enhanced N400 and P600. As a first check to this aim we fitted additional GAMM models using the same syntax as in the manuscript but adding tensors including the available variables of the stimulus that could have influenced the response (that is, quantifier word frequency and semantic relatedness (similarity) across the cue and the upcoming target word). For transparency, the full list of stimuli is reported in Table 14S. To make the models computationally manageable, we had to run two different models, one featuring semantic relatedness, and the other featuring word frequency. Here we report the models' syntax for the word frequency model:

```
Ampl ~ Cond # main effect
+ s(word_frequency, k = 5) # main effect
+ s(Time, by = Cond) # main effect
+ s(norm_log_power) # main effect
+ ti(Time, norm_log_power, k = c(20, 20), by = Cond) # two-way interaction between time and pre-stimulus power, depending on condition
+ ti(Time, word_frequency, k = c(20, 5), by = Cond) # two-way interaction between time and word frequency, depending on condition
+ ti(Time, norm_log_power, word_frequency, k = c(20, 20, 5), by = Cond) # three-way interaction of interest (time, pre-stimulus power and frequency, depending on condition)
+ s(Time, ID, bs = "fs", m = 1) # random effect
```

```
+ s(Time, cond, bs = "fs", m = 1) # random effect
+ s(Event, bs = "re") # random effect
+ s(Time, Event, bs = "re") # random effect
```

The one for semantic relatedness was identical, except for the variable of interest. We run the models on some representative electrodes (CP1, CP2, P3 and P4).

Results for the three-way interaction of interest between time, pre-stimulus alpha power and word frequency were not significant (see Table11S for the p values), meaning that controlling for word frequency did not significantly modulate the association between pre-stimulus power and ERPs amplitude levels. On the other hand, the two-way interaction between time and pre-stimulus alpha power remained mostly significant (see Table 12S), and as can be seen from pictures 12 to 15S, the modulation of ERP amplitudes by means of pre-stimulus alpha power was very similar to that obtained in the original models, reported in the paper. We can therefore conclude that word frequency dies not explain the trial-by-trial variability in ERP amplitudes.

Table 11S. Significance of the three way interactions between time, pre-stimulus alpha power and word frequency depending on condition. Significant results are marked with *.

Electrodes	Interaction: time*power* frequency, incongruent condition	Interaction: time*power* frequency, congruent condition
CP1	0.71	0.29
CP2	0.35	0.89
P3	0.60	0.30
P4	0.75	0.68

Table 12S. Significance of the two way interactions between time and pre-stimulus alpha power, depending on condition. SIgnificant results are marked with *.

Electrodes	Interaction: time*power, incongruent condition	Interaction: time*power, congruent condition
CP1	0.04 *	0.01*
CP2	< 0.001*	0.27
P3	< 0.001*	0.05*
P4	< 0.001*	0.13

Figure 12S.

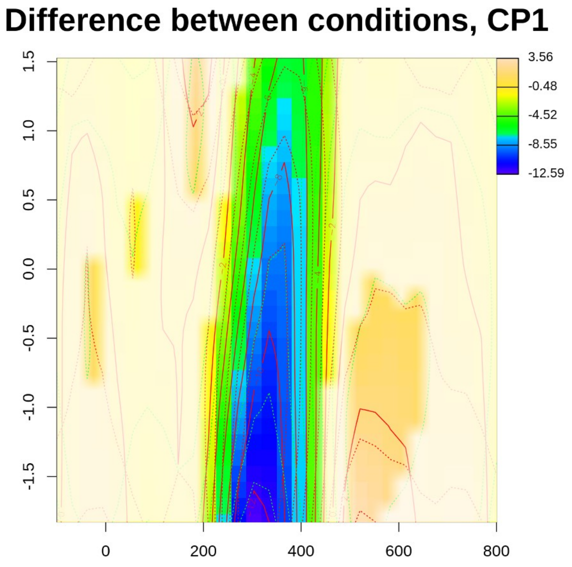


Figure 13S.

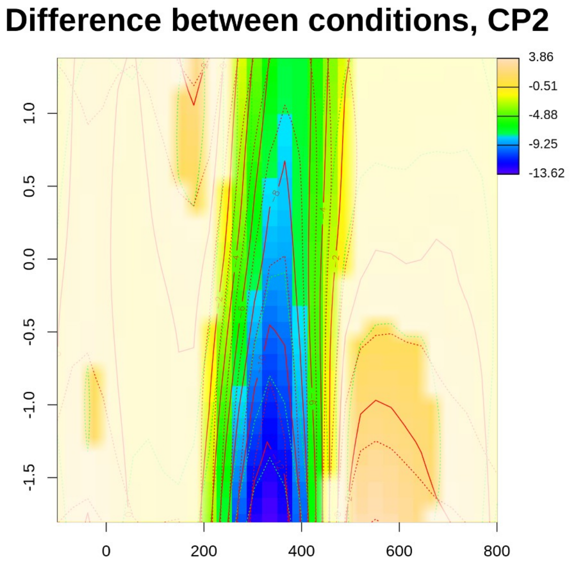
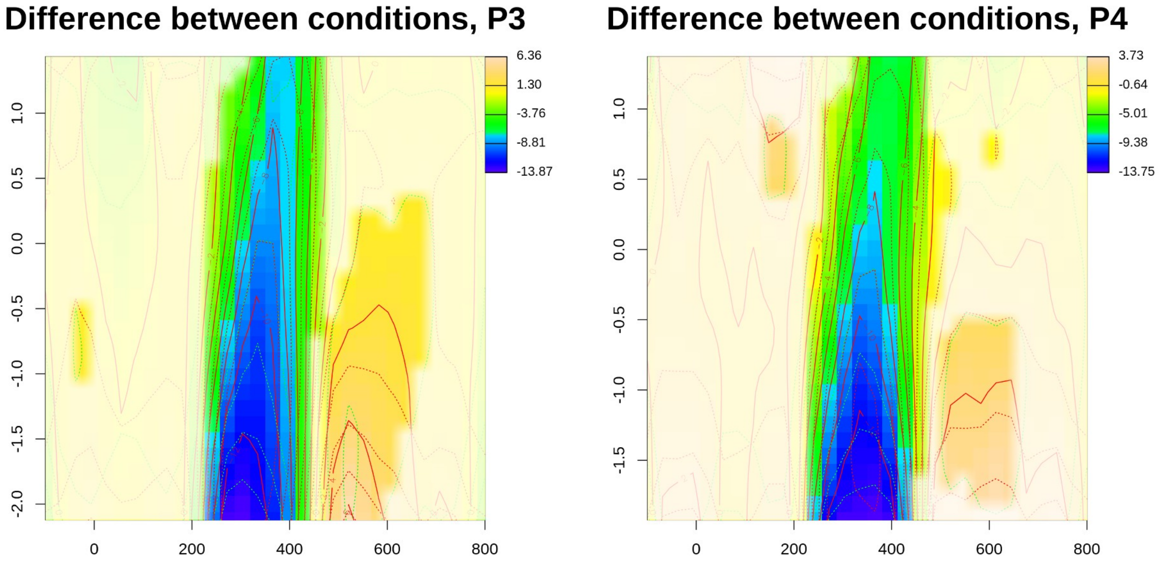


Figure 14S.

Figure 15S.


Difference between conditions, P3

Difference between conditions, P4

Figures 12 to 15S. Predicted amplitude (tensor surface) differences between incongruent and congruent conditions for the two-way interaction (time and pre-stimulus alpha power). The figure depicts the N400 and P600 effects depending on the pre-stimulus alpha level. Time is represented on the x-axis and pre-stimulus power on the y-axis. Colored areas reflect significant predicted ERP amplitude differences between congruent and incongruent responses. Shades of yellow and red indicate increasingly positive differences, while shades of green and blue indicate increasingly negative differences. Green shades indicate that the difference tends to zero. White areas indicate regions where the confidence intervals (95%) around the predicted surface included zero.

The relatedness models that included three way interactions were killed for computational memory overload, therefore we had to simplify the model syntax by removing the three way interaction (syntax reported below):

```
Ampl ~ Cond # main effect

+ s(relatedness) # main effect

+ s(Time, by = Cond) # main effect

+ s(norm_log_power) # main effect
```

```
+ ti(Time, norm_log_power, k = c(20, 20), by = Cond) # two-way interaction between time and pre-stimulus power, depending on condition
+ ti(Time, relatedness, k = c(20), by = Cond) # two-way interaction between time and semantic relatedness, depending on condition
+ s(Time, ID, bs = "fs", m = 1) # random effect
+ s(Time, cond, bs = "fs", m = 1) # random effect
+ s(Event, bs = "re") # random effect
+ s(Time, Event, bs = "re") # random effect
```

Results for the two-way interaction of interest between time and semantic relatedness depending on condition is significant (see Table 13S), meaning that in these electrodes there is a significant effect of semantic relatedness on the ERPs.

Table 13S. Significance of the two way interaction between time and semantic relatedness depending on condition. Significant results are marked with *.

Electrodes	Interaction: time*semantic relatedness, incongruent condition	Interaction: time*semantic relatedness, congruent condition
CP1	0.02*	0.01*
CP2	0.02*	0.04*
P3	0.09	0.03*
P4	0.01*	0.01*

However, upon visual inspection of the effect (see figures 16 to 19S), we can see that different levels of semantic relatedness are not predictive of different ERP amplitudes (i.e., ERP amplitude is independent from the level of semantic relatedness).

Figure 16S.

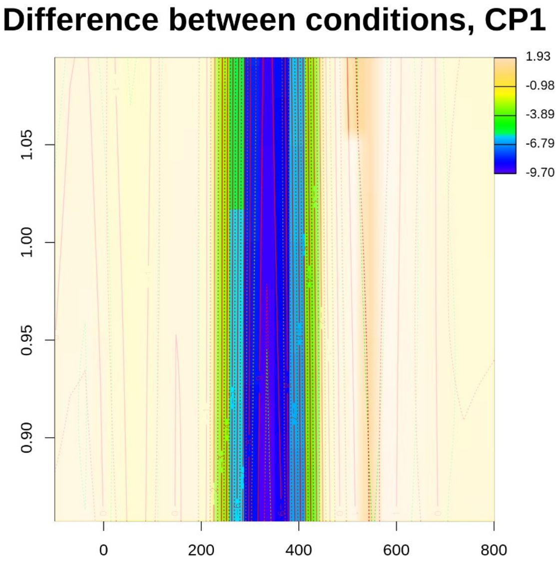


Figure 17S.

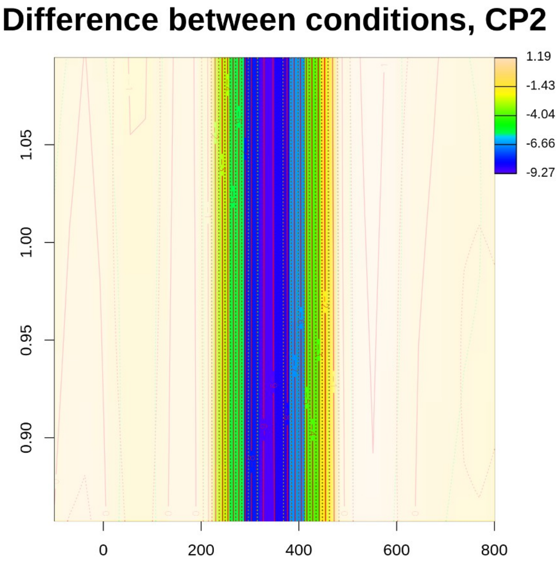


Figure 18S.

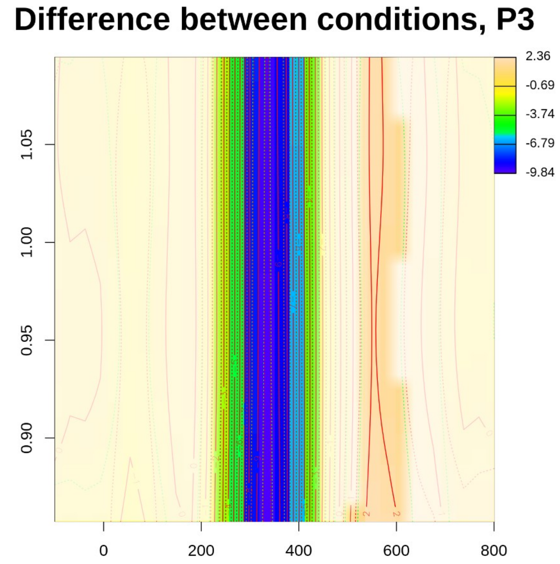
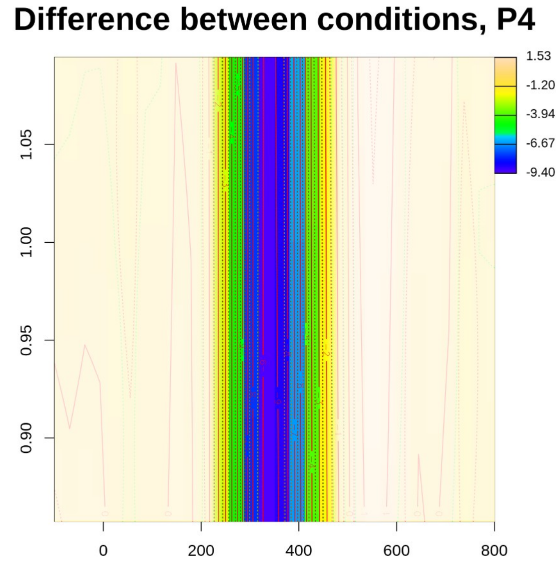


Figure 19S.



Figures 16 to 19S. Predicted amplitude (tensor surface) differences between incongruent and congruent conditions for the two way interaction (time and semantic relatedness). The figure depicts the N400 and P600 effects depending on semantic relatedness level. Time is represented on the x-axis and relatedness on the y-axis.

In other words, semantic relatedness is associated with the presence of a N400 and/or late posterior positivity/P600 effect depending on electrodes, but it does not explain the trial-by-trial variations of ERP amplitudes, since amplitude is constant regardless of the semantic relatedness level. On the other hand, figures 20 to 23S show the effect of the interaction between time and pre-stimulus alpha power depending on condition; these images are comparable to those reported for the original models. Given these results, we can conclude that even when controlling for semantic relatedness, the effect of pre-stimulus power is still a better predictor of the trial-by-trial variations of ERP amplitudes.

Figure 20S.

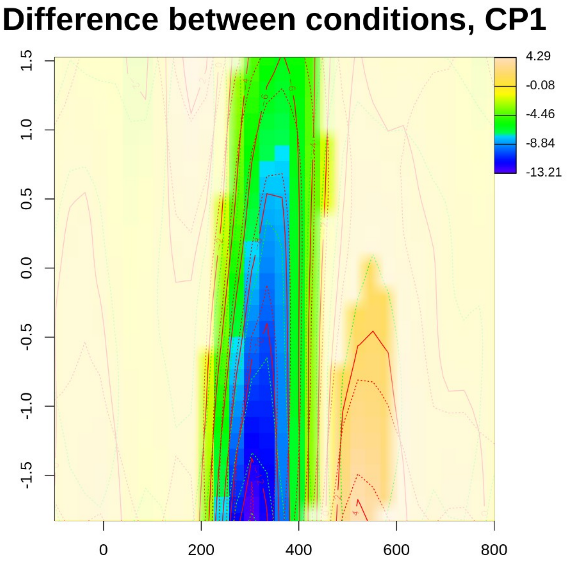


Figure 21S.

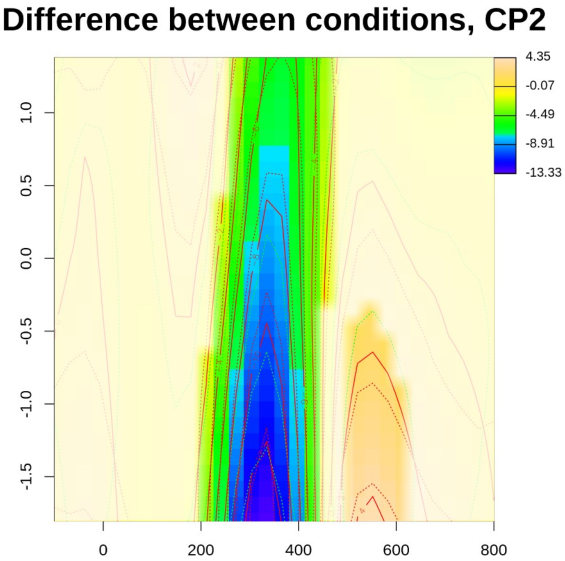
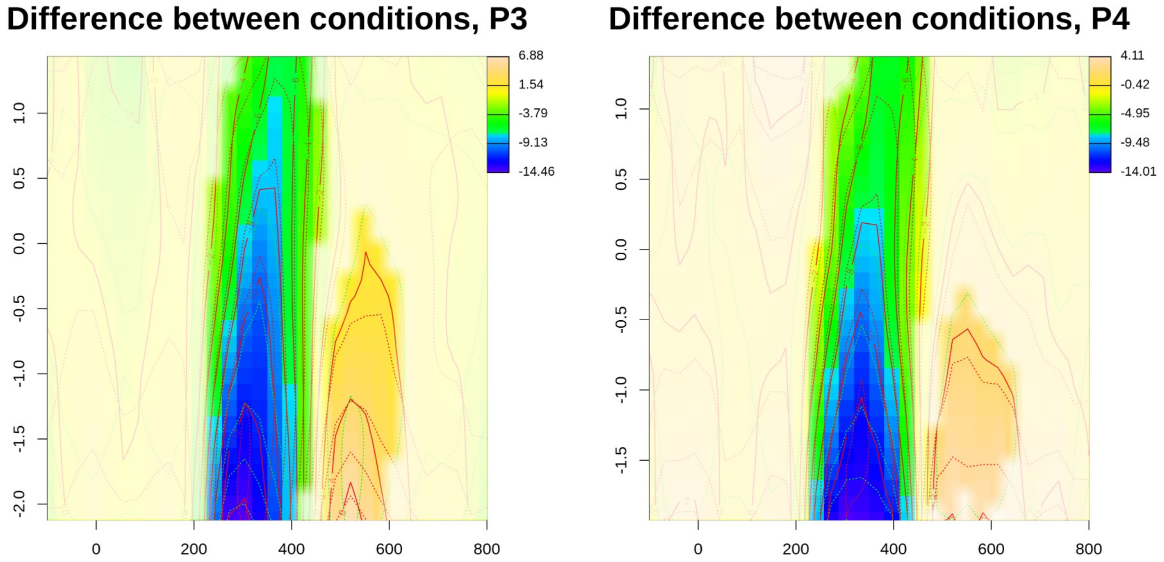


Figure 22S.

Figure 23S.


Difference between conditions, P3

Difference between conditions, P4

Figures 20 to 23S. Predicted amplitude (tensor surface) differences between incongruent and congruent conditions for the two-way interaction (time and pre-stimulus alpha power) of the semantic relatedness model. The figure depicts the N400 and P600 effects depending on the pre-stimulus alpha level. Time is represented on the x-axis and pre-stimulus power on the y-axis.

In sum, results of the models with the additional available variables show that these variables did not change the conclusions that can be derived from the original models. Even if these additional models show that some stimuli variables have no influence on the main results of the experiment, it is important to underline that other cue-target word pairs' properties could be of more relevance, e.g. cue reliability, i.e., how much a figure is associated with a specific noun. This issue posits two additional problems: i) it is not clear how to evaluate cue reliability in the context of this specific experimental task; ii) given the quasi-experimental design of the study, it is never possible to exclude the presence of further additional missing confounding variables, other than cue reliability.

To further substantiate the claim that the observed values of alpha are related to random fluctuations we run a simulation using the following logic:

1) If the fluctuations of alpha power are associated with some properties of the stimuli, then some stimuli should be systematically associated with lower pre-stimulus alpha power as compared to others. This is because, for some reason, cue validity or any other missed confound would be associated more often with a lower (or higher) value of pre-stimulus power;
2) We can rank, separately for each participant, the stimuli associated with the lowest prestimulus power, considering the N lowest stimuli;
3) We can run a Monte Carlo Simulation in which we simulate 1000 times a dataset analogous to the real one, but randomly assigning the rank of all the stimuli. For each simulation we store the value of a stimulus that is ranked more often in the range of N more extreme stimuli (e.g. for lower alpha ranking, suppose that the stimulus with highest occurrence rate is a stimulus ranked 97 times as in the lower alpha range, than the number "97" is stored as max observed value). We can build a distribution of all maximum observed values under random ranking and this would act as null-distribution for our statistical test (note that this simulation is strongly inspired to Mass Univariate Statistics, Groppe et al., 2011 and Cluster Based Permutation approach, Maris & Oostenveld, 2008).

In other words, the simulation we built answers the following question: "what is the expected probability, under the null hypothesis, of observing that some specific items tend to be ranked as the lowest?". Albeit with limitations, this is an empirical test of the "randomness" of the observed data distribution, regardless of the potential confound taken into account.

Figure 24S reports the empirical distribution observed on our data, counting how often the item was among the 20 items with the lowest alpha power values. The horizontal line depicts the threshold of 95° percentiles of distribution of values as obtained in the simulation.

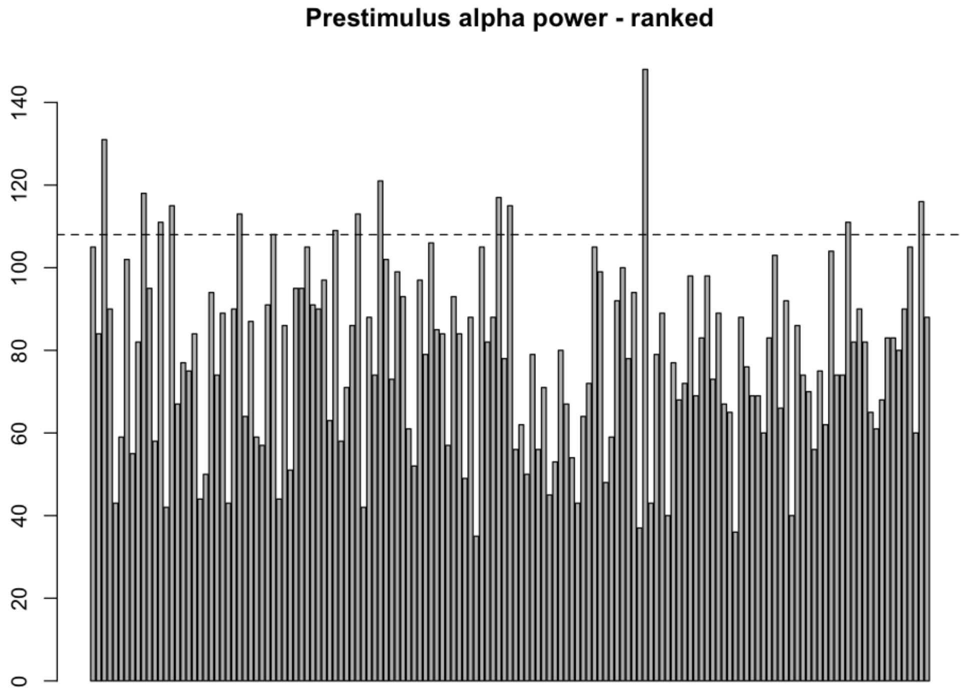


Figure 24S. Results of observed stimulus ranking (of prestimulus alpha power) as compared to random ranking. Each bar represents a single stimulus (cue-target word pair) and how many times it appears ranked as one within the 20 pairs with the lowest alpha (as calculated within participants). The dashed line indicates the results of the simulation of random ranking as described above. In particular, when randomly ranking the stimuli 95% of the stimuli were ranked within the 20-th lowest no more than 108 times (hence the threshold is 108). All bars exceeding this threshold indicate that they tended too often (as compared to random ranking) to be among the ones with the lowest prestimulus alpha.

Results from this simulation show that 8.7% of the items showed values exceeding the threshold obtained by the simulation. This result is close to those expected when the null hypothesis is true (in which, by definition, 5% of values exceeded this threshold). As additional control, we re-fit the same model reported in the manuscript excluding those items, that is excluding the items which tended too often to be among the ones with the lowest alpha variables. Results of this trimmed model are consistent with the original models including all the stimuli (see Figures 25 and 26S), suggesting that the observed effect is not related to those specific stimuli that tended to show the lowest prestimulus alpha.

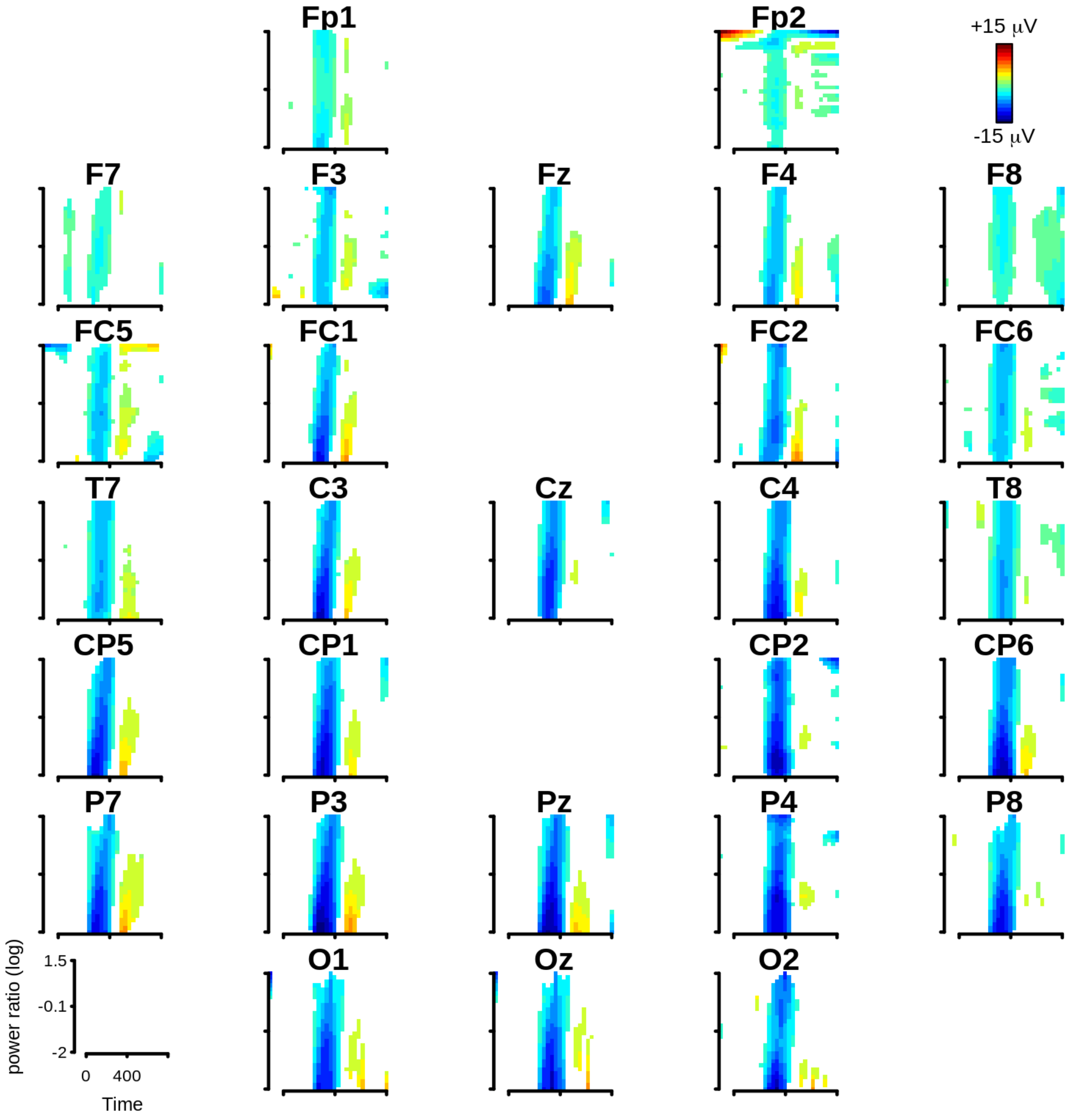


Figure 25S. Differences between incongruent and congruent tensor surfaces for the interaction between time, pre-stimulus alpha power, and condition from the trimmed models.

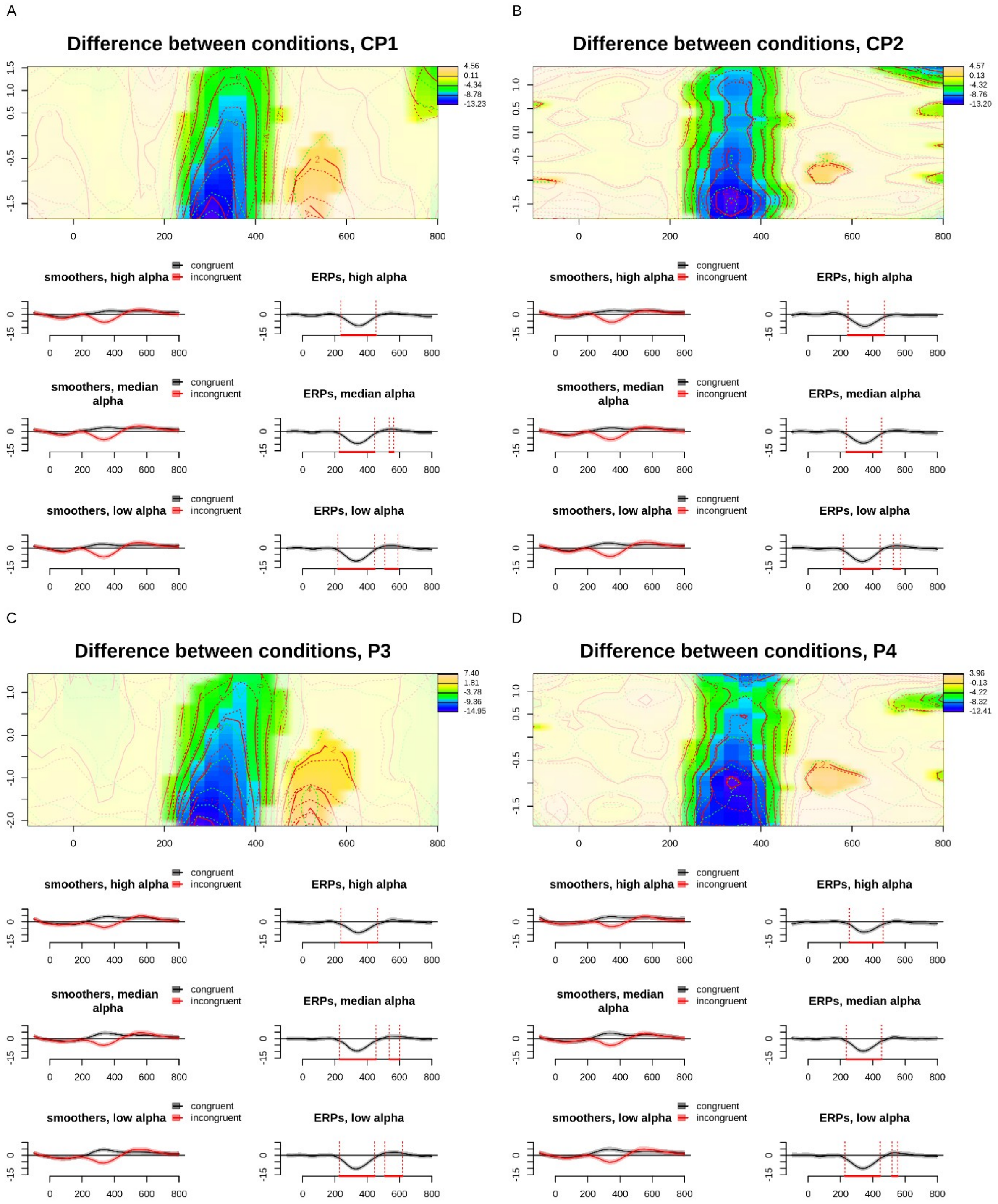


Figure 26S. Highlights for specific electrodes from the trimmed models. Results are analogous to those reported in the main text.

Additional references

Bornkessel, I. D., Fiebach, C. J., Friederici, A. D., & Schlesewsky, M. (2004). “Capacity” reconsidered: Interindividual differences in language comprehension and individual alpha frequency. *Experimental Psychology*, *51*(4), 279–289. https://doi.org/10.1027/1618-3169.51.4.279

Grandy, T. H., Werkle-Bergner, M., Chicherio, C., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2013). Individual alpha peak frequency is related to latent factors of general cognitive abilities. *NeuroImage*, *79*, 10–18. https://doi.org/10.1016/j.neuroimage.2013.04.059

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. Psychophysiology, 48(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Howard, C. J., Arnold, C. P. A., & Belmonte, M. K. (2017). Slower resting alpha frequency is associated with superior localisation of moving targets. *Brain and Cognition*, *117*(January), 97–107. https://doi.org/10.1016/j.bandc.2017.06.008

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods, 164(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Nalaye, H., Cross, Z. R., Schlesewsky, M., & Bornkessel-schlesewsky, I. (2022). Electrophysiological indices of individual differences in adult language learning. *BioRxiv*, 1–43.

Ociepka, M., Kałamała, P., & Chuderski, A. (2022). High individual alpha frequency brains run fast, but it does not make them smart. *Intelligence*, *92*(February). https://doi.org/10.1016/j.intell.2022.101644

Ramsay, I. S., Lynn, P., Schermitzler, B., & Sponheim, S. (2021). Individual alpha peak frequency is slower in schizophrenia and related to deficits in visual perception and cognition. *Scientific Reports*, *11*(1), 1–9. https://doi.org/10.1038/s41598-021-97303-6

Samaha, J., & Postle, B. R. (2015). The Speed of Alpha-Band Oscillations Predicts the Temporal Resolution of Visual Perception. Current Biology, 25(22), 2985–2990. https://doi.org/10.1016/j.cub.2015.10.007

Study 3

1. Materials: psycholinguistic details of the stimuli

Stimulus properties were defined by asking a series of participants (who did not perform the MEG task) questions about the stimuli. First we enrolled a total of 17 participants to investigate properties about sentence completion of experimental sentences (age: mean = 29, SD = 6.29, range = 22-46; education: mean = 16, SD = 2.74, range = 13-21; 12 Females). These participants were presented with the same sentences of the experimental items, with the exceptions of the last word, which was not presented (e.g. “That lawyer is a ____”). They were asked to complete the sentences with the first word that came up in their mind. From this data collection, two measures were computed: Cloze Probability was calculated as the percentage of people that completed the sentence with the same word used in the experimental stimuli (e.g. in reference to the examples as in Fig. 1, “shark”, in the sentence, “that lawyer is a ____ ”). Entropy was calculated to express a measure of variability in responses given by the participants. It was calculated as

$$H=\sum_{\square}^{\square} p(x)\cdot\log p(x)$$

Where H is the entropy associated with a stimulus and p(x) was the probability of observing a given response in the tested sample. Entropy is high when there are many equiprobable responses, while it is low when there is just one single response or one dominant response. In a few words, while Cloze Probability captures the probability of predicting correctly the actual target given the preceding context, Entropy estimates how variable is the set of candidates that is predicted given the preceding context.

For Familiarity a total of 39 participants were enrolled (age: mean = 27.05, SD = 4.54, range = 20-43; education: mean = 16.69, SD = 2.44, range = 11-21, 20 Females). They were asked to rate, on a 5-point likert scale, the degree of familiarity of the stimuli, from 1 (not familiar at all) to 5 (completely familiar).

We also added three corpus based measures: log word Frequency, as collected from the itWAC corpus (Baroni, Bernardini, Ferraresi, & Zanchetta, 2009), target word length, calculated as number of letters, and similarity between the first noun of the sentence and the target word (e.g., between "lawyer" and "shark" in the example "that lawyer is a shark"), expressed as cosine distance calculated from WEISS 3-words window context (Marelli, 2017).

We investigated differences between the two stimulus types (Metaphors and Literal) by means of independent t-tests and by equivalence tests in all the above mentioned measures. While t-tests investigate differences in the stimuli, equivalence tests investigate whether groups can be considered as significantly equal. Metaphoric stimuli had lower familiarity and lower Cloze Probability as compared to Literal ones. Stimuli also showed a difference in Entropy (with higher entropy for metaphoric contexts) and a higher semantic similarity between the two nouns. Overall these differences are all expected and intrinsically related to the kind of stimuli, as we employed non lexicalized metaphors that almost by definition are less common than their literal counterparts. Note that for frequency and length no test is reported because the target stimuli on which these two variables were calculated were exactly the same for metaphoric and literal sentences.

Details on stimuli and comparisons are reported in Table 1S.

Variable	**df**	**t**	**Cohen's d**	**P value**	**Equivalence p**	**epsilon**	**Mean Lit (SD)**	**Mean Met (SD)**
Cloze Probability	160	4.2	0.66	< 0.001	< 0.001	0.1	0.05 (0.09)	0.01 (0.03)
Entropy	160	-1.6	-0.24	0.120	< 0.001	0.3	3.68 (0.34)	3.75 (0.27)
Familiarity	160	6.1	0.96	< 0.001	0.021	1.000	3.4 (0.62)	2.65 (0.91)

Target Length	160	-	-	1.000	-	-	7.17 (2.12)	7.17 (2.12)
Log Frequency	160	-	-	-	-	-	8.47 (1.45)	8.47 (1.45)
Similarity	160	9	1.4	< 0.001	1.000	0.1	0.33 (0.14)	0.15 (0.1)

Table 1S. The table reports the mean variables associated with experimental stimuli (Metaphors and Literal sentences). Each row shows a measure. Columns report: the name of the variable, the degrees of freedom of the t-tests for the comparison Metaphor vs Literal, the t-value for this test, the p-value for this test, the equivalence p-value (which is the p-value stating if groups can be considered significantly equal), the epsilon value, which is required for the equivalence test, mean and SD for the Metaphor and literal sentences.

2. Statistical analyses

R syntax for models with ROI activity on TPJ as dependent variable:

```
ROI_activity ~ cond + s(Time, by=cond) + s(TPJ_log_pre_power) # main effects

+ ti(Time, TPJ_log_pre_power, k=c(20, 20), by=cond) # tensor with main interaction

+ s(Time, subj_ID, bs=”fs”, m=1) # nonlinear random smoother with individual differences in time course

+ s(Time, target, bs=”fs”, m=1) # nonlinear random smoother with differences in time course for each target word

+ s(Event, bs=”re”) # random intercept for each trial time series

+ s(Time, Event, bs=”re”) # random slope for each trial time series
```

The R syntax for models with Language ROI activity as dependent variable is reported here:

```
ROI_activity ~ cond + s(Time, by=cond) + s(ROI_log_pre_power) +
s(rTPJ_log_pre_power) + s(lTPJ_log_pre_power) # main effects

+ ti(Time, ROI_log_pre_power, k=c(20, 20), by=cond) + ti(Time,
rTPJ_log_pre_power, k=c(20, 20), by=cond) + ti(Time,
lTPJ_log_pre_power, k=c(20, 20), by=cond) # tensors with main
interactions

+ s(Time, subj_ID, bs=”fs”, m=1) # nonlinear random smoother with
individual differences in time course

+ s(Time, target, bs=”fs”, m=1) # nonlinear random smoother with
differences in time course for each target word

+ s(Event, bs=”re”) # random intercept for each trial time series

+ s(Time, Event, bs=”re”) # random slope for each trial time series
```

3. Results

3.1. Whole-brain cluster-based permutation t-test

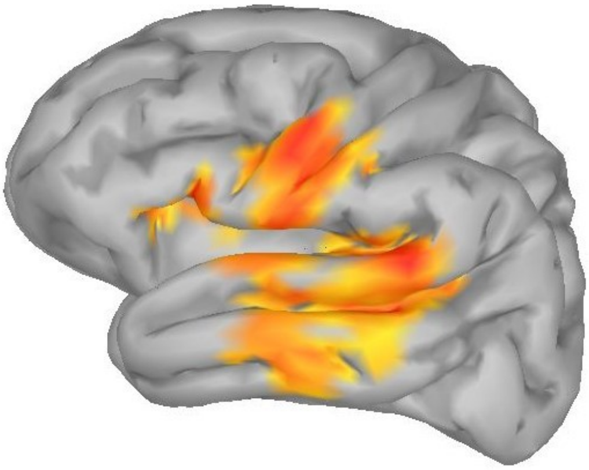


Figure 1s. Results of the cluster-based permutation t-test on whole-brain activations. There was a significant positive activation difference between the metaphorical and literal conditions. Activations were higher for Metaphors as compared to Literal expressions.

3.2. Same-area interactions for the language-related ROIs models

Table 2s reports the significance of the interaction between time, pre-stimulus alpha power, and literal vs metaphorical condition for the language-related ROIs. Within the literal condition, same-area pre-stimulus alpha power only predicted the post-stimulus activation within the MTG. Within the metaphoric condition instead, same-area pre-stimulus alpha power only predicted the post-stimulus activation within Broca's area.

Table 2S. Significance of interactions between time, power, and condition within the language-related areas. Significant results are marked with *.

	literal condition			metaphorical condition		
	Broca	MTG	STG	Broca	MTG	STG
Interaction: time, same area power	0.38	<0.001*	0.4	0.01*	0.4	0.07

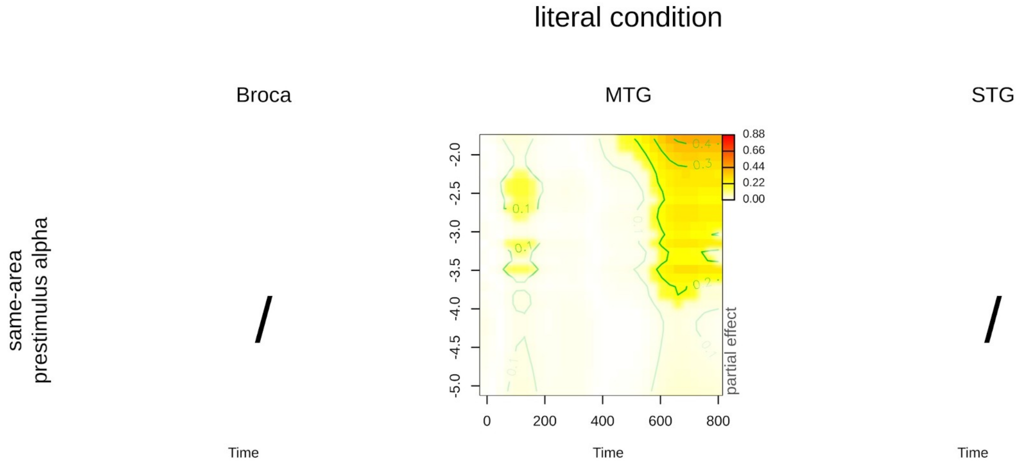


Figure 2s. 3-D plots of the main interaction between time and pre-stimulus alpha power for the literal condition, for significant effects. Time is represented on the x-axis, pre-stimulus alpha power level on the y-axis, and activation is color-coded. Colored blots represent plot regions where the interaction is significant, while white areas indicate regions where the confidence intervals (95% CI) around the predicted surface included zero, i.e., the interaction is not significant. Increasingly dark shades indicate greater activation.

As we can see from Figure 1s, in the literal condition, within the MTG, pre-stimulus alpha power from this area itself is also associated with late activations (after 500 ms).

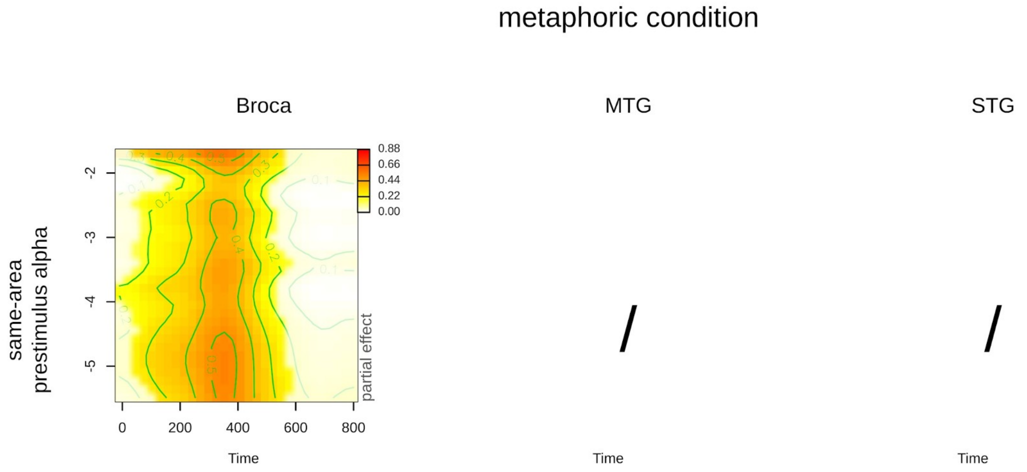


Figure 3s. 3-D plots of the main interaction between time and pre-stimulus alpha power for the metaphorical condition, for significant effects. Time is represented on the x-axis, pre-stimulus alpha power level on the y-axis, and activation is color-coded. Colored blots represent plot regions where the interaction is significant, while white areas indicate regions where the confidence intervals (95% CI) around the predicted surface included zero, i.e., the interaction is not significant. Increasingly dark shades indicate greater activation.

As we can see from Figure 2s, within Broca's area, pre-stimulus alpha power from the area itself is associated with early activation (until about 400 ms), resembling the effect in the literal condition (see main text), but more extended in time.

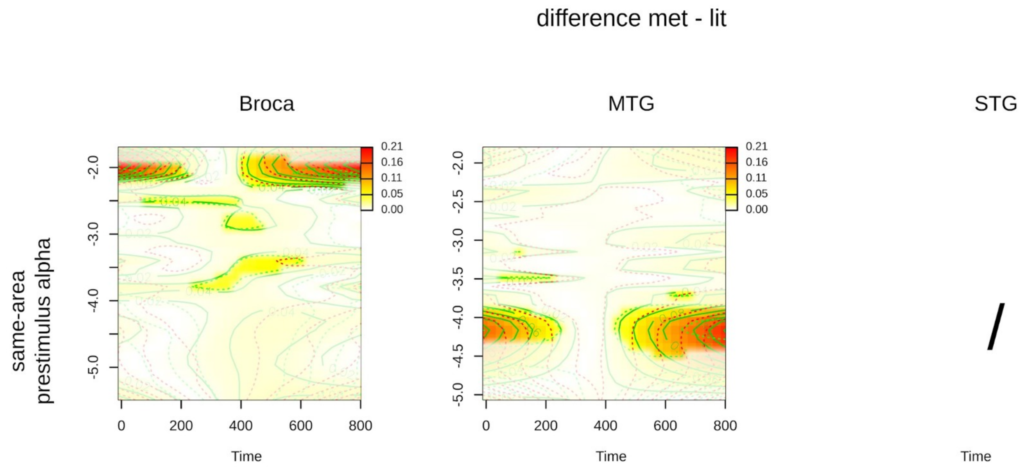


Figure 4s. Predicted activation (tensor surface) differences between metaphorical and literal conditions for the main interactions of interest (time, alpha band power, and conditions). Time is represented on the x-axis, pre-stimulus alpha power level on the y-axis, and predicted activation difference is color-coded. Colored blots represent plot regions where the difference is significant, while white areas indicate regions where the confidence intervals (95% CI) around the predicted activation difference included zero, i.e., the difference is not significant. Increasingly dark shades indicate greater activation differences.

As can be seen in Figure 3s, activation differences in Broca's area and in the MTG are also predicted by the pre-stimulus alpha power recorded from these same areas. In both ROIs, significant differences are present from 0 to 200 msec and after 400 msec; however, such differences are associated with high pre-stimulus alpha power in Broca's area and with lower pre-stimulus power in the MTG.

3. Control analyses

Since the pre-stimulus interval we chose included the indefinite article ("a/an") and we did not explicitly control for possible anticipatory processes happening at the level of the article in the task design, we performed additional analyses. First of all, we have to point out that, given the quasi-experimental nature of the design, even in case we establish that anticipatory processes are taking place already at the level of the article, we could not rule out the explanation that a further additional confound associated with the stimuli could explain the fluctuations in alpha power taking place in the pre-stimulus interval. Please note that this limitation is not specific for our study but is shared by any quasi-experimental study in language that investigates properties that are not manipulated (e.g., language studies investigating difference between two types of stimuli in a lexical decision task).

To further substantiate our claims we thus followed a different way of reasoning, that involved simulations, which are now reported in the supplemental data. The logic used was the following:

1) If the fluctuations of alpha power are associated with anticipatory processes relative to the determinant, then some stimuli should be associated more often with lower (or higher) pre-stimulus alpha power as compared to others. This is because those stimuli would be associated for some reason (i.e., congruency of the determinant with the participants' hypotheses or any other missed confound) with a lower (or higher) value of prestimulus power.
2) We can rank, separately for each participant, the stimuli associated with the lowest or highest prestimulus power, considering the N lowest or N highest stimuli (and using different values of N).
3) We can run a Monte Carlo Simulation in which we simulate 1000 times a dataset analogous to the real one, but randomly assigning the rank of all the stimuli. For each simulation we store the value of stimulus that more often is ranked in the range of N more extreme stimuli (e.g. for lower alpha ranking, suppose that the stimulus with highest occurrence rate is a stimulus ranked 97 times as in the lower alpha range, than the number "97" is stored as max observed value). We can build a distribution of all max observed values under random ranking and this would act as null-distribution for our statistical test (Note that this simulation is strongly inspired to Mass Univariate Statistics, Groppe et al., 2011 and Cluster Based Permutation approach, Maris & Oostenveld, 2007).

In other words, the simulation we built answers the following question: “what is the expected probability, under the null hypothesis, of observing that some specific items tend to be ranked as the lowest?”. Albeit with limitations, this is an empirical test of the “randomness” of the observed data distribution, regardless of the potential confound taken into account.

The figure below reports the empirical distribution observed on our data, counting how often the item was among the 20 items with the lowest alpha power values. The horizontal line depict the threshold, as obtained in the simulation.

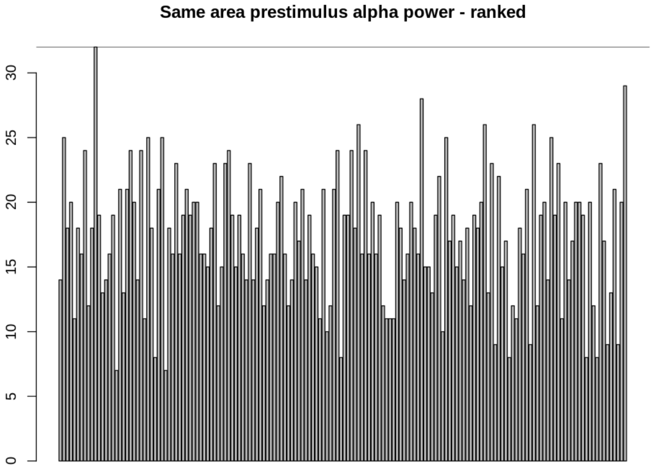


Figure 5s. Results of observed stimulus ranking (of prestimulus alpha power) as compared to random ranking. Each bar represents a single stimulus (article-target word pair) and how many times it appears ranked as one within the 20 pairs with the lowest alpha (as calculated within participants). The line indicates the results of the simulation of random ranking as described above. In particular, when randomly ranking the stimuli 95% of the stimuli were ranked within the 20-th lowest no more than 36 times (hence the threshold is 36). All bars exceeding this threshold indicate that they tended too often (as compared to random ranking) to be among the ones with the lowest prestimulus alpha.

Other published works

The following studies are (however tangentially) related to what has been outlined in the present work.

- Lago, S., Bevilacqua, F., Stabile, M. R., Scarpazza, C., Bambini, V., & Arcara, G. (2022). Case report: Pragmatic impairment in multiple sclerosis after worsening of clinical symptoms. Frontiers in Psychology, 13, 1028814. https://doi.org/10.3389/fpsyg.2022.1028814
    - Study 1 reviewed several domains of cognitive impairments resulting from several neurological pathologies. However, in the review the language domain was not covered adequately. This case report presents an interesting account of decline of pragmatic abilities, that can be interpreted, in light of the findings outlined in the present dissertation, as a possible global impairment of predictive processing. Future investigations are needed to elucidate the specific neural and cognitive processes that might be involved in the decline of pragmatic abilities.
- Sulpizio, S., Arcara, G., Lago, S., Marelli, M., & Amenta, S. (2022). Very early and late form-to-meaning computations during visual word recognition as revealed by electrophysiology. Cortex, 157, 167–193. https://doi.org/10.1016/j.cortex.2022.07.016
    - This study allowed me to implement and perfectionate the statistical analyses technique used in Studies 2 and 3.

Response to Reviewers

Reviewer 1

The present thesis represents a nice contribution to the literature on predictive processes and the relationship between pre-stimulus neurobiological correlates and post-stimulus brain activity. It provides a nice overview on the role of the temporo-parietal junctions during language processing and contributes to further deepen our understanding on the interaction between TPJ and other language-related brain areas. Study 1 provides a review of the studies on the role of right TPJ in neurological patients and proposes a general functional interpretation of these brain areas within a domain-general predictive framework. Study 2 describes an EEG experiment investigating the relationship between pre-stimulus alpha and post-target evoked responses (N400/P600) in a picture matching design. Study 3 is an MEG experiment on literate and figurative language processing aimed at testing the link between pre-stimulus alpha in TPJs and temporal dynamics of post-target neural activity in other language-related areas.

These studies represent a nice contribution to the literature on language processing and predictive analysis. The designs of the studies are sound and well justified. The introduction of each study is well structured. I highly appreciated the thoughtful explanation of each of statistical and methodological choice. I also want to highlight that the candidate shows a strong expertise in a wide range of electrophysiological methods, ranging from EEG to MEG. Overall, I think that the candidate can be admitted to the final exam.

I have some minor suggestions that I hope can help widening the theoretical introduction, smoothing the between-studies transitions, and deepening the understanding of scientific results presented here (especially for Study 3).

- Introduction, page 11: when talking about alpha and its functional interpretation in the cognitive domain (especially language), I think it would be good to mention the “Gating and pulsed inhibition hypothesis” (Jensen & Mazaheri, 2010; Frontiers) and recent evidence suggesting a relevant role of alpha in reading (Jensen, Pan, Frisson, & Wang, 2021, TICS; Pan, Popov, Frisson, Jensen, 2023, Plos Biol).
  Answer: Thank you for the suggestion. I edited the text accordingly (p.11).
- Introduction, page 47: when taking about possible functionally different sub-areas of rTPJ it would be good to know if this distinction can be supported by cytoarchitectonic differences (something that has been suggested for the visual word form area as well; e.g., Weiner KS, Barnett MA, Lorenz S, et al. 2017, Cereb Cortex).
  Answer: Thank you for the suggestion. I edited the text accordingly (p.48).

- Study 1, page 48: can you give an estimation about how often rTPJ is solely damaged vs how often it is damaged together with other brain areas that are in the territory of the middle cerebral artery? In other words, how easy is to disentangle the functional interpretation of rTPJ from the surrounding areas in clinical studies?
  Answer: As reported in Table S3, of the 54 studies included in the review, most (33) reported either a direct lesion of rTPJ (28) or a direct lesion in addition to the disconnection of rTPJ from surrounding areas (5), in relation to the behavioral deficit observed. Among these 33 studies, only 5 reported that rTPJ was solely damaged, while the large majority reported lesions of other areas in addition to lesions to rTPJ. This probably indicates that other frontal, parietal, and temporal regions can frequently show abnormalities in neurological pathologies, while small, selective lesions circumscribed to rTPJ might occur very rarely. This consideration highlights the impossibility of disentangling the effects of lesions solely to rTPJ from lesions to rTPJ *and* other brain areas with the only aid of clinical, observational studies; however, it is also in line with the observation, outlined in the Introduction, that rTPJ is part of many neural networks subserving different cognitive functions. For this reason, it is beyond the scope of this study to disentangle the effects of a lesion circumscribed to rTPJ from the effects of a lesion that also includes the surrounding areas, because we argue that brain regions are indeed interconnected. Therefore, even if a lesion only involves one area, such a lesion might entail functional abnormalities in other regions of the network too. Our goal was rather to give a more holistic picture of the role of the area of interest, without focusing only on lesions solely circumscribed to rTPJ, but taking into account all the cases when this region was involved in behavioral and cognitive manifestations of neurological diseases. The fact that our results demonstrate how rTPJ is frequently involved in many different complex behaviors and cognitive domains, either alone or with other brain areas, supports once again the hypothesis that this area might be part of a domain-general predictive network.
- Transition between Study 1 and Study 2: it is not clear what is the logical transition connecting the two studies here. Are there studies supporting the idea that the neural sources of alpha, N400 or P600 might include TPJs? How is TPJ related to the role of pre-stimulus alpha and why should we suspect a connection there?
- *Answer:* Thank you for the suggestion. I added an interim summary (p. 75)
- Study 2, Page 91-92: very nice description of the statistical details of the study and very clear justification of the methodological choices.
  Answer: Thank you, your feedback is greatly appreciated.

- Study 2, page 111: to further support the conclusion here ("the N400 amplitude was less sensitive to pre-stimulus power level" [as compare to the P600 amplitude]), it would be good to add a quantification of the modulatory effect of pre-stimulus alpha power on N400 and P600 amplitude, respectively.
  Answer: Thank you for the suggestion. I edited the text accordingly (p.112).
- Transition between Study 2 and Study 3: again, it is not clear what is the logical transition connecting the two studies here. Is there any study suggesting that the neural sources of pre-stimulus alpha power include TPJ? This can be also added and further elaborated on pages 133-134.
  Answer: Thank you for the suggestion. I added another interim summary (p. 126).
- Study 3, materials: the experimental design comes a little bit out of the blue here and it should be anticipated more in the last part of the introduction to familiarize the reader with the experimental conditions and the specific related hypotheses.
  Answer: Thank you for the suggestion. I edited the text accordingly (p. 136).
- Study 3, page 143: "The words included in this interval ("is a/an") are not content words and they are consistent throughout the sentence stimuli, so their effect on sentence comprehension should be constant across all sentences." There can be specific anticipatory processes happening already at the level of the determinant, which might influence pre-stimulus alpha power (DeLong, Urbach & Kutas, 2005; Nature). How are these aspects controlled in this design?
  Answer: We thank the reviewer for this comment, which we consider to be really relevant. We did not explicitly control for possible anticipatory in the task design. However, we addressed this point in several ways. First of all, given the quasi-experimental nature of the design, even in case we establish that anticipatory processes are taking place already at the level of the determinant, we could not rule out the explanation that a further additional confound associated with the stimuli could explain the fluctuations in alpha power taking place in the pre-stimulus interval. Please note that this limitation is not specific for our study but is shared by any quasi-experimental study in language that investigates properties that are not manipulated (e.g., language studies investigating difference between two types of stimuli in a lexical decision task).

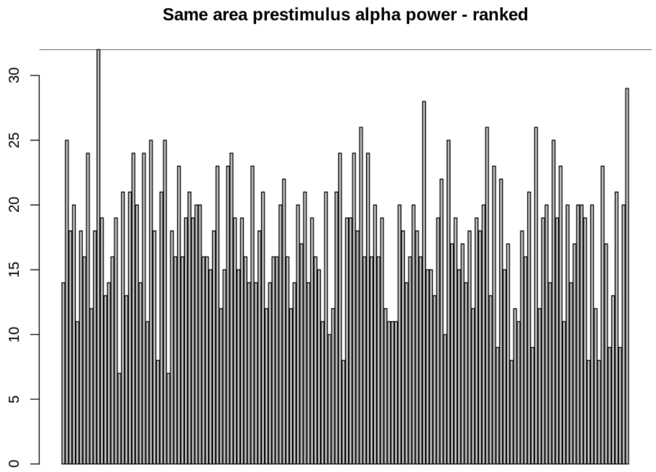
  To further substantiate our claims we thus followed a different way of reasoning, that involved simulations, which are now reported in the supplemental data. Note that the same strategy was used to solve a similar issue for Study 2 (See Supplementary Materials, p. 249 and following).

  The logic used was the following:

1) If the fluctuations of alpha power are associated with anticipatory processes relative to the determinant, then some stimuli should be associated more often with lower (or higher) pre-stimulus alpha power as compared to others. This is because those stimuli would be associated for some reason (i.e., congruency of the determinant with the participants' hypotheses or any other missed confound) with a lower (or higher) value of prestimulus power.
2) We can rank, separately for each participant, the stimuli associated with the lowest or highest prestimulus power, considering the N lowest or N highest stimuli (and using different values of N).
3) We can run a Monte Carlo Simulation in which we simulate 1000 times a dataset analogous to the real one, but randomly assigning the rank of all the stimuli. For each simulation we store the value of stimulus that more often is ranked in the range of N more extreme stimuli (e.g. for lower alpha ranking, suppose that the stimulus with highest occurrence rate is a stimulus ranked 97 times as in the lower alpha range, than the number "97" is stored as max observed value). We can build a distribution of all max observed values under random ranking and this would act as null-distribution for our statistical test (Note that this simulation is strongly inspired to Mass Univariate Statistics, Groppe et al., 2011 and Cluster Based Permutation approach, Maris & Oostenveld, 2007).

In other words, the simulation we built answers the following question: "what is the expected probability, under the null hypothesis, of observing that some specific items tend to be ranked as the lowest?". Albeit with limitations, this is an empirical test of the "randomness" of the observed data distribution, regardless of the potential confound taken into account.

The figure below reports the empirical distribution observed on our data, counting how often the item was among the 20 items with the lowest alpha power values. The horizontal line depict the threshold, as obtained in the simulation.



Results for this simulation are now reported in supplemental data, Figure 5s of Study 3, and show that none of the items showed values exceeding the threshold obtained by the simulation. This result suggests that no item is associated systematically with a particularly lower value of alpha.

These analyses and results are now reported also in the Supplementary Materials (p. 261 and following) and mentioned in the Methods section (p. 143 and following) and in the Discussion (p. 162).

- Study 3, page 147-148: do these complex models converge?

  Answer: yes, they do, and I edited the text on page 148.
- Study 3, page 149-150: were there any significant differences between conditions in the behavioral results?

  Answer: yes, please see p. 149 for a comment
- Study 3, page 150: was there any TPJ difference in the whole brain results?

  Answer: no, significant differences did not encompass any of the TPJs. This information is now reported on p. 150
- Study 3, page 150: I highly encourage the authors to report full statistical details and not only p-values. E.g., Table 2 and following Tables should report beta values, standard errors and t (or z) values.

  Answer: thank you for the suggestion. I edited the tables accordingly.
- Study 3, page 151: was there any relationship between TPJ pre-stimulus alpha and the pre-stimulus alpha activity of language-related brain areas? Are we sure that the effect of TPJ on language-related areas starts after the presentation of the target?

  Answer: I will answer the first part of this question here, while the answer to the second part (Are we sure that the effect of TPJ on language-related areas starts after the presentation of the target?) is provided together with the answer to the next question because I think they are related.

First of all, I checked whether there was a relationship between TPJ pre-stimulus alpha and the pre-stimulus alpha activity of language-related brain areas by means of GAMMs, with the formula:

ROI_prestim_power ~ s(lTPJ_prestim_power) + s(rTPJ_prestim_power)

Results are reported in Table 1 and they show that rTPJ pre-stimulus alpha power levels, in particular, have an effect on the pre-stimulus alpha in all the language-related areas, while lTPJ pre-stimulus alpha power only modulates STG pre-stimulus power. To my opinion, however, this is an expected result and it is in line with the theoretical assumption stated throughout the dissertation. Prediction is an ongoing and pervasive process that, in the case of the present experiment, takes place throughout the sentence. Predictive processing also presumably involves a distributed network of areas, encompassing the TPJs (see Study 1), whose role is context- and network-dependent. Therefore, the TPJs and the other areas involved in predictive computation within a particular cognitive domain unavoidably influence one another at different time points.

Table 1. Beta coefficients and *p* values of the main effects of left and right TPJ pre-stimulus power on the language-related ROIs pre-stimulus power. Significant results are marked with *.

	Broca pre-stimulus power		MTG pre-stimulus power		STG pre-stimulus power	
	Beta	p value	Beta	p value	Beta	p value
lTPJ pre-stimulus power	2.042	0.096	1.957	0.16	6.603	<0. 001*
rTPJ pre-stimulus power	9.265	<0. 001*	6.919	0.001*	27.429	<0. 001*

- Study 3, page 154: Figure 6 left upper corner plot seems to show a difference that starts before 0 ms. Did the authors check what happens before the target presentation? Is there a baseline issue? How do the authors explain such an early effect?

  Answer: In addition to these questions, I will also answer part of the question above: Are we sure that the effect of TPJ on language-related areas starts after the presentation of the target?

As happens with the time-frequency decomposition in EEG/MEG analyses where results show smearing effects and are therefore excluded from the interpretation, the edges in our graphs might show unreliable and spurious results since the datapoints in these areas are less dense than in other areas of the graph, such as in the center. For these reasons, we did not draw any conclusions regarding effects that are too close to the edges, such as those apparently starting before 0 ms. Also, we did not investigate what happened before the target stimulus in terms of ROI activation or of the relationship between pre-stimulus alpha power and pre-stimulus activation level. However, we have to consider prediction as a continuous and ongoing process, so it it plausible that the TPJs exert a continuous influence on other brain areas throughout the sentence, as we have demonstrated in the previous point, and that the activation in the language-related areas could be influenced by the TPJs alpha level during a time interval spanning the whole sentence.

In this context, studying whether pre-stimulus alpha could modulate the activation in the same pre-stimulus time window could not be the best strategy, since we are investigating the same phenomenon with two different methodologies and we would probably end up finding an influence of one variable on the other simply because we are looking at the same time interval. On the other hand, a more sensitive approach would be a "sliding window" one, i.e. to take TPJ alpha power from the pre-stimulus interval before each word and examine whether and how it modulates the post-stimulus activation in the time interval spanning from the presentation of each word to the starting point of the next one, thereby investigating whether effects really start before 0 ms.

However, this goal was beyond the scope of the present study, but it is a really nice starting point in future studies, and I included these considerations on p. 162.

- Page 176: there is an incomplete sentence at the end of the first paragraph.

Answer: thank you, I edited the text accordingly.

Reviewer 2

General Evaluation:

The thesis explores an intriguing and highly relevant topic in the field of cognitive neuroscience and is well-written and organized. The investigation into the predictive role of TPJ in both domain-general and domain-specific cognition is of great significance and holds the potential for meaningful contributions to the field. The inclusion of diverse studies in the thesis, encompassing a systematic review of patient studies, as well as empirical investigations utilizing neuroimaging techniques (EEG/MEG) in healthy volunteers, underscores the comprehensive nature of the research, which leverages a wide array of methodological approaches. The fact that two of the three studies have been published in excellent peer-reviewed journals Neuroscience and Biobehavioral Reviews, Psychophysiology) speaks to the high quality of the work and its recognition within the academic community. I only have a few suggestions to enhance the overall quality of the thesis.

General Introduction:
The thesis starts with a clear introduction. However, it would be beneficial to provide explicit definitions for the terms 'domain-general' and 'domain-specific' cognition to ensure that readers have a clear understanding of these concepts from the outset. This will help in establishing a solid foundation for the subsequent studies.
Answer: Thank you for the suggestion. I added a clarification at the end of p. 6.
Study 3:

- In the third study, there is an opportunity to streamline the introduction to make it more concise, focusing on the core objectives and hypotheses. This will help maintain reader engagement and clarity.

Answer: Thank you for your suggestion. The Introduction is now shorter.

- Regarding the methodology in Study 3, a clarification about the rationale behind the order of data preprocessing steps, particularly why downsampling was performed before filtering, would be valuable for readers seeking a deeper understanding of the methodology.

Answer: Thank you. I corrected the text (p. 142)

- A important point to consider is that the TPJ, as shown in Figure 1 (General Introduction) and used in Chapter 3, may not correspond precisely to regions described in the Introduction ("bilateral ventral portions of the inferior parietal lobule (the supramarginal and angular gyri") and posterior sections of the superior temporal gyrus) and to the specific cytoarchitectonic areas studied in the literature (|PFop, PFt, PF, PFm, and PFcm, located in he supramarginal gyrus, and the PGp and PGa, located in the angular gyrus"). Clarifying whether this mismatch is a limitation of the study would add depth to the discussion. Moreover, providing information about which cytoarchitectonic areas the TPJ corresponds to, based on the current research, and discussing any potential implications or limitations associated with this discrepancy would be beneficial.

Answer: Thank you for your suggestion. I added these considerations both in the Methods section (p. 142) and in the Discussion (p. 157 and 163).

- On page 147, moving complex formulae to an appendix could improve the readability of the thesis and to maintain focus on the main text.

Answer: Thank you for your suggestion. The models' structures are now reported in the Supplementary Materials, p. 258 and following.

- On page 149, please add a '%' symbol after "96.04".

Answer: Thank you for noticing, I corrected the typo.

- Considering potential confounding factors, it would be helpful to discuss whether the results in Study 3 could be influenced by variations in task difficulty.

Answer: Thank you for your suggestion. I added some consideration on p. 161.

- Lastly, on page 126, a typographical error should be corrected to "pre-activating."

Answer: Thank you for noticing, I corrected the typo.

General Conclusion:

There is a notable omission in the discussion of how the results from the different studies fit into an overarching framework. Highlighting how the findings from each study collectively contribute to the broader understanding of TPJ's predictive role in domain-specific and domain-general cognition.

Answer: Thank you for your suggestion. I added some considerations (together with those suggested in the note below) in p. 176/177.

Moreover, emphasizing the extent to which the results from different studies can be compared, considering variations in study populations (patients vs. healthy volunteers) and methodologies (brain lesions vs. MEG vs. EEG) would further enrich the conclusion. In summary, Sara Lago's thesis is well-written and addresses an important topic. The suggested revisions aim to enhance the clarity, depth, and cohesion of the research presented throughout the thesis.

Answer: Thank you very much for your positive feedback.