UNIVERSITAT DE
BARCELONA

Neural signatures of semantic anticipation in sentence comprehension

Patricia León Cabrera

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Cognition and Brain
Plasticity Unit

UNIVERSITAT DE
BARCELONA

Neural signatures of semantic anticipation in sentence comprehension

Patricia León Cabrera

Cognition and Brain Plasticity Unit
Department of Cognition, Development and Educational Psychology
University of Barcelona

Doctoral program in Brain, Cognition and Behaviour

Supervisor:

Dr. Antoni Rodríguez-Fornells

Co-supervisor:

Dr. Joaquín Morís

Table of contents

List of abbreviations

CNV	Contingent negative variation
EEG	Electroencephalography
ERP	Event-related potential
HC	High constraint
IFG	Inferior frontal gyrus
LC	Low constraint
MCI	Mild cognitive impairment
MEG	Magnetoencephalography
MTG	Middle temporal gyrus
NS	Non-semantic
PD	Parkinson's disease
PD-MCI	Parkinson's disease with mild cognitive impairment
PD-NC	Parkinson's disease with normal cognition
pMTG	Posterior medial temporal gyrus
RP	Readiness potential
RPSV	Rapid visual serial presentation
SCP	Slow cortical potential
SNP	Slow negative potential
SPN	Stimulus-preceding negativity
TF	Time-frequency
WM	Working memory

1 Abstract

It has been proposed that the human brain is a proactive processor. Rather than passively receiving information, it is seen as continuously attempting to predict what will happen next. In the domain of language, evidence supports that people predict different linguistic aspects of the sentences they read or hear, such as the meaning of upcoming words. However, the brain mechanisms that sustain the formulation of these predictions and the subsequent generation of semantic expectations during comprehension remain largely unknown.

The current dissertation investigated neural correlates associated with anticipatory processing during sentence comprehension throughout four electrophysiological (EEG) studies. To this end, a novel experimental paradigm was designed, which allowed to capture two theoretically distinct temporal phases in from the lens of predictive processing: the anticipatory and the processing phase of (un)predicted words. The first three studies focused on exploring the mechanisms involved in the anticipatory phase in healthy young adult population. The first study described signatures of anticipatory processing associated with semantic prediction in speech comprehension. Specifically, in the interval between the context and the final word of a sentence, a sustained negative potential developed, with a larger amplitude at increasing levels of semantic expectancy. The second study replicated this anticipatory index in reading comprehension, establishing generality across input modalities. Furthermore, it demonstrated that differences emerged earlier and progressively over the course of sentence processing. Lastly, the third study revealed that contextually expected words were also preceded by a transient alpha power decrease, common in both modalities of comprehension. Based on their psychophysiological features, the observed neural signatures are consistent with the online recruitment of anticipatory mechanisms tied to semantic prediction during sentence comprehension.

Finally, the fourth study evaluated the status of the previously described brain signatures in adults with Parkinson's disease (PD) (with dopaminergic compensation), given that this condition entails cognitive deficits that might negatively affect the proper use of sentential contexts. Relative to the control group, the group of PD patients exhibited normal correlates of semantic anticipation and semantic processing. On the other hand, semantic processing was altered in a subgroup of PD patients with mild cognitive impairment (PD-MCI), compared to patients without MCI. Specifically, there was a significant prolongation in the semantic processing of words that did not match the contextual expectation. Lastly, for all PD patients, worse verbal fluency scores correlated with alterations in semantic anticipation and processing, suggesting that, in PD, deficits in mechanisms that rely on brain networks in the temporal lobe might hinder predictive language processing during sentence comprehension.

Taken together, the current doctoral dissertation reinforces the notion that individuals mobilize cognitive resources during sentence comprehension to get ahead of information, and significantly contributes to the available corpus of empirical knowledge in dilucidating neural correlates associated with this phenomenon in healthy and pathological conditions. In light of this evidence, a theoretical approach is proposed with the goal of inspiring and guiding future research in the field of predictive language comprehension.

2 Resumen

Se ha propuesto que el cerebro humano es un procesador proactivo, que lejos de recibir información de forma pasiva, está continuamente tratando de predecir qué va a ocurrir a continuación. En el ámbito del lenguaje, existe evidencia empírica de que las personas predicen información lingüística de las frases que leen o escuchan, como, por ejemplo, el significado de palabras que todavía no han percibido. Sin embargo, se desconocen en gran medida los mecanismos cerebrales que subyacen la generación de estas predicciones y la consiguiente generación de expectativas durante la comprensión del lenguaje.

En esta tesis doctoral se presentan cuatro estudios de electroencefalografía (EEG) sobre correlatos neurales asociados al procesamiento anticipatorio durante la comprensión de frases. Se desarrolló un paradigma experimental que permitió capturar dos fases teóricamente diferenciables en el marco de la predicción, la fase anticipatoria y la fase de procesamiento de la información predicha. Los tres primeros estudios de la tesis se centraron en explorar los mecanismos implicados en la fase anticipatoria de la predicción en población no patológica. El primer estudio describió, por primera vez, correlatos de procesamiento anticipatorio asociados a la predicción lingüística durante la comprensión auditiva de frases. En el intervalo entre el contexto y la palabra final de una frase, se observó la aparición de una negatividad sostenida, con una amplitud mayor cuanto más fuertemente esperada a nivel semántico era la palabra final. El segundo estudio replicó la observación de este índice anticipatorio en la modalidad de comprensión escrita, demostrando su carácter amodal. Además, permitió determinar que las diferencias en la actividad neural en función de la expectativa semántica emergían pronto y aumentaban progresivamente a lo largo del procesamiento de las frases. Por último, el tercer estudio reveló que las palabras contextualmente esperadas estaban también precedidas por una desincronización neuronal transiente en alfa, común en ambas modalidades de comprensión. Por sus características psicofisiológicas, los correlatos observados son consistentes con mecanismos anticipatorios vinculados a la predicción de aspectos semánticos durante la comprensión de frases.

Finalmente, el cuarto estudio evaluó el estado de los mecanismos previamente descritos en adultos con enfermedad de Parkinson (PD) (con compensación dopaminérgica), dado que esta patología cursa con déficits cognitivos que afectan el uso apropiado contextos oracionales. El grupo de pacientes con PD exhibió correlatos normales de anticipación y procesamiento semántico comparados con el grupo control. Por otra parte, se encontró que el procesamiento semántico estaba afectado en un subgrupo de pacientes con PD y deterioro cognitivo leve (DCL), respecto a pacientes sin DCL. En concreto, se observó una prolongación significativa del procesamiento semántico de aquellas palabras que no encajaban con la expectativa del contexto oracional. Por último, en toda la muestra de PD, una peor fluidez verbal correlacionó con alteraciones en anticipación y procesamiento semánticos, sugiriendo que déficits en mecanismos dependientes de circuitos temporales en pacientes con PD podrían mermar el procesamiento predictivo durante la comprensión de frases.

La presente tesis doctoral refuerza la noción de que las personas movilizan recursos cognitivos durante la comprensión del lenguaje con el fin de anticipar información, y contribuye sustancialmente a nivel empírico, mostrando correlatos neurales relacionados con este fenómeno cognitivo en condiciones normales y patológicas. A la luz de esta evidencia, se propone también una aproximación teórica con el fin de inspirar y guiar futuras investigaciones en el ámbito del procesamiento predictivo del lenguaje.

Introduction

3 Introduction

Many of the events that happen in our environment do not come out of the blue. As we are driving along a road, the transition of the traffic light from green to amber allows us to predict that it will turn red next and that we must stop. In the last decade, the ability to extract regularities and patterns from the world to predict events has been situated at the core of brain function. Not only do we predict when there are blatant cues, such as a traffic light, but our brains are relentlessly attempting to get ahead of the future (Bar, 2007; Bar, 2009). Originally inspired by the study of human perception, this view of the brain as a predictive organ has been progressively extended to a variety of cognitive domains, including action, attention, or learning (Clark, 2013), promising to offer a parsimonious and neurobiologically plausible explanation of how the brain manages to optimally navigate an ever-changing, highly uncertain and noisy world (Friston, 2005, 2010). Indeed, to some, the brain is, essentially, a prediction machine (Clark, 2013).

Joining this trend, several models of language comprehension have also embraced prediction (Kuperberg, 2013; Kuperberg & Jaeger, 2016; Kutas, DeLong, & Smith, 2011; Pickering & Gambi, 2018). Comprehending language appears to be very simple for most readers and listeners and, yet, it is extraordinarily complex, requiring the successful coordination of multiple cognitive processes under high time constraints. It has been hypothesized that predictive processing may account, at least to some extent, for the easiness and efficiency with which comprehension is achieved.

Considering the centrality of language in our life, uncovering the neural mechanisms that enable and support the generation of predictions in sentence comprehension is an important research endeavour. In the following sections, I will first define what is meant by predictive language processing. Then, I will provide an overview of the available empirical evidence in support of prediction in sentence comprehension, with a focus on electrophysiological findings. Most of this evidence has focused on the neural substrates of making either correct or wrong predictions about certain words. In turn, there is a paucity of knowledge about the neural mechanisms that underlie the formulation of predictions *before* such words are read or heard. I will later elaborate more on the implications of this void of knowledge and explain how the current dissertation aimed to contribute to filling this gap of knowledge.

3.1 From incremental to predictive language processing

Language comprehension involves two core components: accessing the meaning of individual words in a sentence and combining them to build a global representation of the meaning of that sentence. This latter process of linking current and prior information into a cohesive mental model (Graesser, Singer & Trabasso, 1994; Singer, Graesser & Trabasso, 1994) is known as integration. For many years, the dominant models of language comprehension viewed integration as a “delayed” process that started only after the meanings of all words in the sentence had been accessed (Fodor, 1983; Forster, 1985). This implied that information from the sentence context was not readily used to guide processing – the external, bottom-up input took precedence. This perspective was in part due to the use of discrete and end-state measures, mainly reaction times, which cannot provide very nuanced information about how language processing unfolds over time. Later, the introduction of time-resolved methods, such as eye-tracking, helped unveil a different picture. For instance, when listeners follow spoken instructions to move or touch objects in a visual display, their eye movements to the target objects are closely time-locked to the spoken words that refer to those objects (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995). This strongly suggested that incoming information is processed and combined with the context without further delay (see also, Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

Nowadays, it is largely assumed that comprehension proceeds fast and incrementally. New incoming words are quickly and successively combined and this leads to the progressive construction of a high-level representation of meaning. But, as depicted in **Figure 1**, a language processor that implements prediction goes one step further than that: the information provided by the sentence context starts to exert an influence not immediately, but actually *before* all words are uniquely identified (Kutas & Federmeier, 2011).

There are several ways in which predictive language processing may be more advantageous than ‘waiting’ to incorporate information upon arrival. In fact, prediction may be an optimal solution for some of the main challenges that language comprehension entails. Linguistic input, especially speech, is often noisy. It is not uncommon for other sounds in the environment to partially, or even completely, mask the speech signal, which can sometimes impede the correct identification of words. On top of that, language is fleeting. Speech usually unfolds very rapidly (with rates of about 150 words per minute) which imposes strong time-constraints on comprehension (Christiansen & Chater, 2015). Prediction may help overcome these problems by speeding up the processing of incoming input and guiding interpretation when the input is perceptually ambiguous. Also, prediction may help prevent temporary ambiguities from halting comprehension by allowing to pick up the most likely interpretation of a sentence in advance (Kutas, DeLong & Smith., 2011).

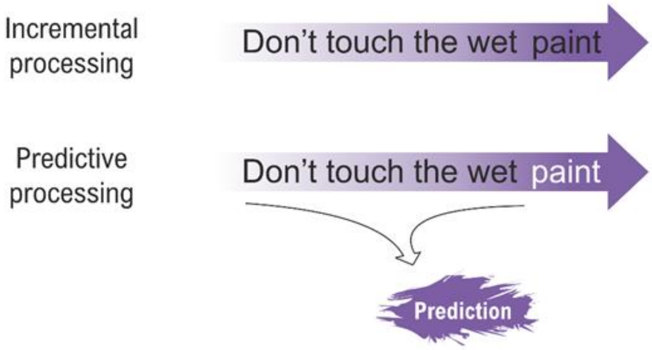

Figure 1. From incremental to predictive language processing. The purple right-pointing arrows represent the incremental construction of sentence meaning. In predictive processing, the unfolding information is used to form predictions about upcoming input, for example, in this case, about the sentence-final word (highlighted in white).

3.1.1 What we mean by prediction

In the domain of language, the term 'prediction' has been defined in a myriad of ways, a heterogeneity that is manifest in the title of a recent influential review, "What do we mean by prediction in language comprehension?" (Kuperberg & Jaeger, 2016).

In the current dissertation, the term *prediction* will be used broadly to refer to *the pre-activation of information, at some representational level, before there is external evidence for it.* Unless it is explicitly stated, the use of the term 'pre-activation' does not imply that a specific word (i.e., lexical representation) is activated (for examples of these cases, see Section 3.1.3.4). Indeed, as will be set forth in the following sections, there is evidence that different aspects of incoming input at different levels of representation can be pre-activated.

How exactly predictions are realized is an open question and this dissertation remains agnostic in this point. However, we assume that the brain holds probabilistic information about language. There are arguably at least two sources from which this information can be drawn (**Figure 2**):

1) *from statistical dependencies between stored representations*, such that, for example, reading the word ‘rain’ be used to predict “wet” based on their semantic relatedness or associativeness, and/or,

2) *from inferred associations between the sentence representation and stored representations*. As discussed in the previous section, excluding some exceptions (e.g., proverbs), there are no pre-stored representations of the meaning of sentences: they are built from anew. The derived representation contains high-level information that goes beyond the mere sum of word-level associations, and that can be used to infer less evident (and even new) affordances. For example, in ‘Don’t touch the wet’, the beforementioned association between the ‘wet’ and ‘rain’ is no longer relevant, but in that context, it can be inferred that the word ‘paint’ may be of relevance in the proximal future.

Based on these sources of information, future events are assigned a probability; if the assigned probability is high, the input is predictable, if the probability is low, the input is unpredictable.

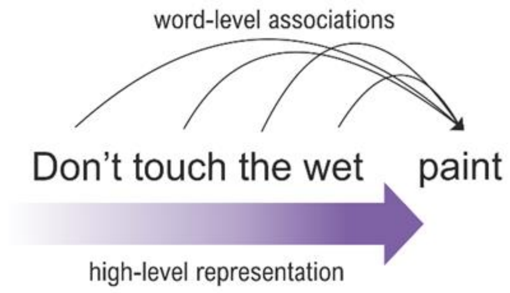

Figure 2. Schematic representation of potential sources of probabilistic information. The fragment “Don’t touch the wet” contains at least two sources of probabilistic information that can be used to predict ‘paint’: 1) pre-existing word-level associations between ‘paint’ and each word in the sentence (black right-pointing arrows), and 2) inferred associations between ‘paint’ and the derived context representation (purple right-pointing arrow).

Furthermore, in addition to the pre-activation of relevant representations, prediction can also entail other forms of preparation for the upcoming input, such as tuning attention or the motor system based on the expected information (Pezzulo, Butz, and Castelfranchi, 2008). We will use the term *anticipation* as an umbrella term to refer to any form of preparation that is put in motion as a result of prediction.

3.1.2 Effects of predictability in language comprehension

There are everyday situations that intuitively hint at the idea that language is predictive. For example, it is not rare to have one's utterances finished by another person, and people can often help children or inexpert language speakers to complete their sentences when they don't know or can't find the words.

On experimental grounds, the predictability of a word can be quantified in terms of cloze-probability – the percentage of individuals that supply that word as a continuation of a sentence in an offline and unspeeded test (Taylor, 1953). For example, the sentence "Don't touch the wet" is continued with the word "paint" by most respondents (1a), while only a few provide "dog" as a continuation (1b):

1a. (High cloze ending) Don't touch the wet <u>paint</u>

1b. (Low cloze ending) Don't touch the wet <u>dog</u>

Thus, in this context, 'paint' is more predictable than 'dog'. The value of the word with the highest cloze-probability (e.g., "paint") defines the semantic constraint of the sentence, that is, how strongly the sentence context drives to one specific word. Highly constraining contexts (2a) lead strongly to a single word – the best completion (e.g., 'paint') –, which has a much higher cloze-probability than any other continuation, whereas low constraining contexts (2b) admit multiple continuations with lower but similar cloze-probabilities:

2a. (High constraint) <u>Don't touch the wet</u> paint

2b. (Low constraint) <u>There was nothing wrong with the</u> car/kid/job

Using the abovementioned cloze-probability measure, numerous studies have shown that readers fixate their gaze less time on words that are predictable and are more likely to skip them (Balota, Pollatsek, & Rayner, 1985; Rayner, Slattery, Drieghe, & Liversedge, 2011; for a review, see Staub, 2015). Furthermore, behavioural tasks that require readers or listeners to make speeded decisions about words find that reaction times are faster for predictable than for unpredictable targets (in naming, Forster, 1981; in lexical and semantic decision tasks, Schwanenflugel & LaCount, 1988; Schwanenflugel & Shoben, 1985; Traxler & Foss, 2000).

3.1.2.1 N400 context effects

More convincing evidence came from event-related potential (ERP) studies. ERPs are voltage fluctuations in the electroencephalographic (EEG) signal recorded from the scalp that are time-locked to a specific event, such as the presentation of a word. Because these voltage fluctuations are tiny, ERPs are usually observed after averaging multiple occurrences of the same event, forming an average ERP waveform (Donchin, Ritter, & McCallum, 1978). ERP waveforms reveal a series of reproducible positive- and negative-going deflections that delineate neural activity with millisecond precision. Crucially, these systematic responses, termed ERP components, reflect specific neural or psychological processes (Kappenman & Luck, 2011).

An ERP component that has been widely used to investigate predictive language processing is the N400 component (Kutas & Hillyard, 1980) (**Figure 3A**). The N400 component was first described by Kutas & Hillyard (1980) in a landmark study in which they compared ERP responses elicited by sentence-final words that were semantically congruent (It was her first day at <u>work)</u> with those that were semantically incongruent (He spread the warm bread with <u>socks</u>). Relative to congruent endings, incongruent endings elicited a larger negative-going deflection that peaked about 400 ms after word onset (**Figure 3B**), and they consequently coined it N400 (Kutas & Federmeier, 2011).

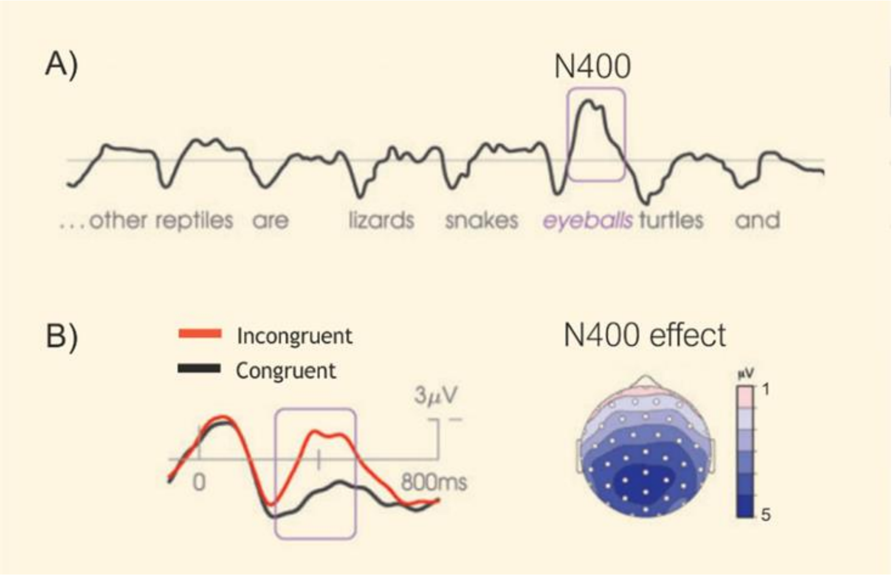

Figure 3. Graphic representation of the N400 response and N400 effect. A) Average ERP to written words, highlighting the larger N400 response to the semantically incongruent word in the sentence (purple frame). **B)** On the left, average ERP comparison of the average ERP at a central midline electrode showing larger N400 amplitudes to incongruent (red) than congruent (black) sentence-final words. On the right, the scalp map of the N400 congruency effect, obtained by subtracting the mean voltages from 300 to 500 ms (purple frame) at every electrode of the incongruent condition from those of the congruent condition. The resulting map exhibits the standard centro-parietal distribution of the N400 effect. Note that, following the convention in ERP research, negative values are plotted upwards in the ERP waveforms. *Adapted from Kutas & Federmeier (2011).*

Although the N400 was first described in response to incongruent words, subsequent studies demonstrated that all words – presented either in isolation or in context, in written or spoken format – produce some degree of N400, including words that are completely congruent (Kutas & Hillyard, 1984). In fact, not only words but virtually any meaningful stimulus elicits an N400-like response, including pictures, faces, comic strips or sounds (Barrett & Rugg, 1989; Barrett & Rugg, 1990; Cohn, 2012; van Petten & Rheinfelder, 1995), which has led to the hypothesis that the N400 is associated with an amodal system implicated in the processing of meaning (Kutas & Federmeier, 2000).

More interestingly, following on their discovery of the N400, Kutas & Hillyard (1984) showed that the amplitude of the N400 is inversely correlated with the cloze-probability (i.e. predictability) of the eliciting word (**Figure 4A**). For instance, a word provided by 80% of participants produced a smaller N400 than a completion provided by 60% of participants. Henceforth, the simpler label 'N400 context effect' will be used to refer to this modulation. This pattern has been reliably replicated in several occasions over the years (Federmeier & Kutas, 1999a; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007a; Wlotko & Federmeier, 2012), with a reported correlation between the amplitude of the N400 and cloze-probability of r = –0.79 (**Figure 4B** and **Figure 4C**) (DeLong, Urbach, & Kutas, 2005).

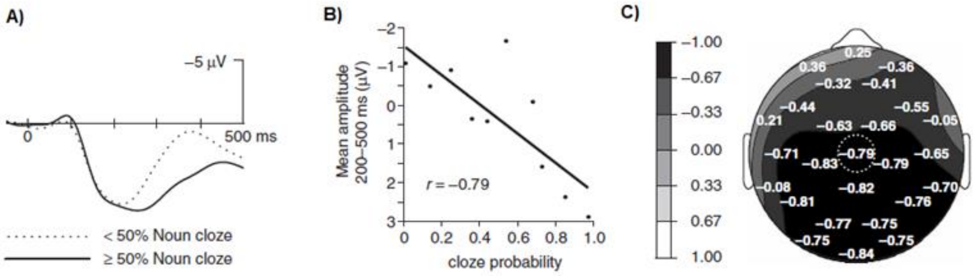

Figure 4. Effect of cloze-probability on the amplitude of the N400. A) On the left, average ERP responses time-locked to sentence-final words (nouns) with higher (>50%) or lower (<50%) cloze-probabilities above or below 50%. The amplitude of the N400 is reduced for words that are more predictable (straight line) compared to less predictable ones (dotted line). Negative is plotted upward. **B)** Correlation between mean N400 (200 to 500 ms post-word onset) and cloze-probability values. The amplitude of the N400 exhibits a significant negative correlation with the cloze-probability of the eliciting word. Best fitting regression line is plotted. C) Topographic map with the r-values for all electrodes. Correlations were maximal over centroparietal sites (darker shading), in line with the usual scalp distribution of the N400 context effect (high minus low cloze-probability). *Adapted from Delong, Urbach & Kutas (2005).*

Moreover, Kutas & Hillyard (1984) also observed that this N400 modulation as a function of cloze-probability was independent of the contextual constraint of the sentence. Final words to high (3a) or low constraining sentences (3b), but with equal cloze-probabilities, elicited indistinguishable N400 responses:

3a. (High constraint; Low cloze) The bill was due at the end of the <u>hour</u>

3b. (Low constraint; Low cloze) He was soothed by the gentle <u>wind</u>

What drove the amplitude variation was the predictability of the word itself, irrespective of the prior context. Note that word predictability can also parsimoniously account for the beforementioned N400 congruency effects (Kutas & Hillyard, 1980), as incongruent words are (often) less predictable than congruent ones. For example, in "*He spread the warm bread with socks*" (1b), the sentence-final word "socks" is not only incongruent but also highly unpredictable. The influence of other lexical variables, such as word frequency, can be similarly explained in terms of predictability – in broad terms, whichever aspect makes a word more likely in a given context seems to ease its processing, yielding smaller N400 responses (Kutas & Federmeier, 2000).

Additional evidence strengthening the link between the N400 and word predictability comes from studies employing word information metrics. Probabilistic language models can be used to estimate different metrics, such as surprisal and entropy, that quantify the amount of information that a word conveys in a sentence (Hale, 2001; Levy, 2008). Word surprisal[1] expresses how unexpected (how 'surprising') the current word is given the previous sentence context, whereas entropy expresses the amount of uncertainty about what will come next in a sentence. Interestingly, word surprisal correlates strongly with the amplitude of the N400 (Frank, Otten, Galli, & Vigliocco, 2015).

The concept of surprisal is closely tied to that of *prediction error*, which is a central notion in current and influential predictive processing frameworks (Clark, 2013; Friston, 2010) (see **Box 1**). In a nutshell, prediction errors result from the comparison between the expected and actual outcome and capture the degree to which a given outcome deviated from the prediction (i.e. how

---

[1] In a way, word surprisal measures can be taken as a more formal specification word predictability than cloze-probability measures, as word surprisal is estimated from probabilistic language models, rather than subjective reports. Yet, one limitation of these information-based measures is that they are largely dependent on syntactic regularities but only minimally sensitive to semantic constraints. Even though it comes with its own shortcomings (as reviewed in Ferreira & Lowder, 2016), the studies in the current dissertation utilize the cloze-probability measure as a proxy of word predictability.

surprising it was). This information is then used as a learning signal to update and adapt the internal probabilistic knowledge, with the goal of improving subsequent predictions. Following this, in recent years, the modulation of the N400 amplitude has been interpreted as reflecting the degree of prediction error (albeit with critical differences in the way prediction error is implemented in each model) (Kuperberg, Brothers, & Wlotko, 2019; Rabovsky, Hansen, & McClelland, 2018; Willems, Frank, Nijhof, Hagoort, & Bosch, 2016; Rabovsky, 2020). For example, in a recent study, Rabovsky and colleagues (2018) successfully simulated N400 context effects with a computational network that treated N400 amplitudes as reflecting *semantic* prediction errors. According to this model, larger N400 amplitudes reflect a larger prediction error, which is used to update the internal probabilistic model from which subsequent predictions about the meaning of the sentence will be derived (see also Rabovsky, 2020).

Box 1. Predictive processing framework

The predictive processing framework (Friston, 2010; Clark, 2013) assumes that the brain is constantly building and updating internal models, based on Bayesian inference, that capture the patterns and statistical regularities of the environment. These probabilistic internal models (also known as forward models) are then used to formulate abstract predictions that propagate from higher to lower-level regions, pre-activating representations of the expected sensory input at different levels in the hierarchy. This framework emphasizes a hierarchical architecture of the brain, but posits that the flow of information across that hierarchy is asymmetrical: the top-down flow carries most information in the form of predictions to 'explain away' the driving sensory signal, whereas only the unexplained portion of the sensory signal (i.e. the 'prediction error') flows upward. The resulting prediction error then serves as a learning signal to update or adapt the internal model and make a better guess in the next iteration of the whole cycle.

Insofar as the N400 is elicited *after* words are presented, N400 context effects would correspond to the processing stage of prediction – they cannot capture the prediction or anticipatory stage, when predictions about *that* word first arise (see **Figure 5**). As we will see next, the post-stimulus nature of N400 component poses several issues when interpreting N400 context effects as unequivocally reflecting prediction.

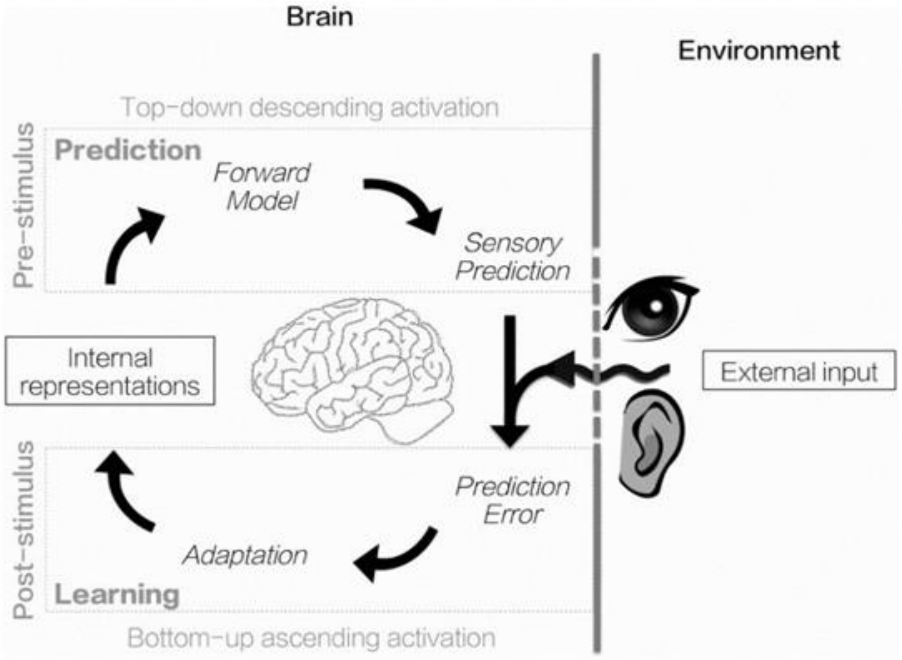

Figure 5. Graphic representation of the different stages involved in prediction according to the predictive processing framework. Predictions are sent top-down from higher to lower level regions to 'explain away' the incoming sensory input (i.e. facilitate its processing). Then, the 'prediction error' – the portion of information that remains unexplained – is sent back and used to tune the internal model to improve future predictions. *Adapted from Molinaro, Monsalve, & Lizarazu (2016)*.

3.1.3 Prediction, or just integration?

Word predictability effects, such as the N400 context effect, can be taken as indirect evidence of prediction: predictable words are easier to process because their representation (or some of their features) have been activated in advance (prediction view). However, given that these effects are examined after the word appears (by contrasting predictable and unpredictable words), it is still possible that differences originate *as* the word is being processed, rather than *before*. Accordingly, predictability effects would be diagnostic of how well the word happens to match (or mismatch) with the previous information, being more or less difficult to integrate with the working context representation (integration view) (**Figure 6**). Note that the latter interpretation does not require any form of pre-activation, instead, the effect would stem from retrospectively comparing the current word with the previous information upon receipt.

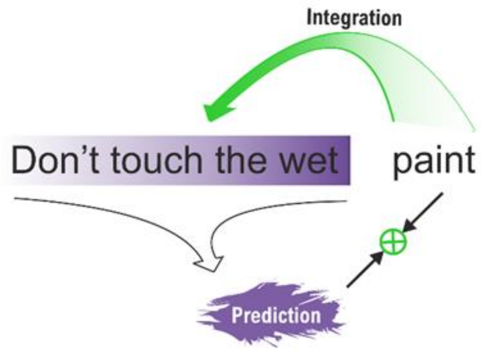

Figure 6. Simplified schematic representation of the prediction and integration views. Incoming words that are more predictable in a given context ('Don't touch the wet paint') are easier to process than less predictable ones (e.g. 'Don't' touch the wet dog') (e.g., Kutas & Hillyard, 1983). This effect of predictability can be explained as a result of prediction, whereby the incoming word is easier to process when it matches the prediction (green equilateral cross), or, alternatively, as a result of integration demands, whereby predictable words are easier to incorporate into the working context than less predictable ones (green left-pointing arrow).

To illustrate this regarding the N400 context effect, let us consider another landmark study by Federmeier & Kutas (1999), in which they compared the N400 response to incongruent endings that shared more (4b) or fewer (4c) semantic features with the expected ending (4a), for example:

They wanted to make the hotel look like a tropical resort.

4a. So along the driveway they planted rows of <u>palms</u>.

4b. So along the driveway they planted rows of <u>pines</u>.

4c. So along the driveway they planted rows of <u>tulips</u>.

Although both PINES and TULIPS are equally anomalous endings, the N400 was smaller for PINES (**Figure 7**). The key aspect is that PINES has a higher featural overlap with the expected word PALMS (within-category, i.e. both are trees) relative to TULIPS (between-category, i.e. it is a flower, not a tree). Under the prediction view, PINES benefited from the prediction of "tree-like" semantic features associated with the expected word PALMS. Alternatively, under the integration view, the semantic features need not be pre-activated in advance. Instead, the sentence context would become relevant only after "pines" is processed and the strong contextual support for "tree-like" features facilitates its incorporation.

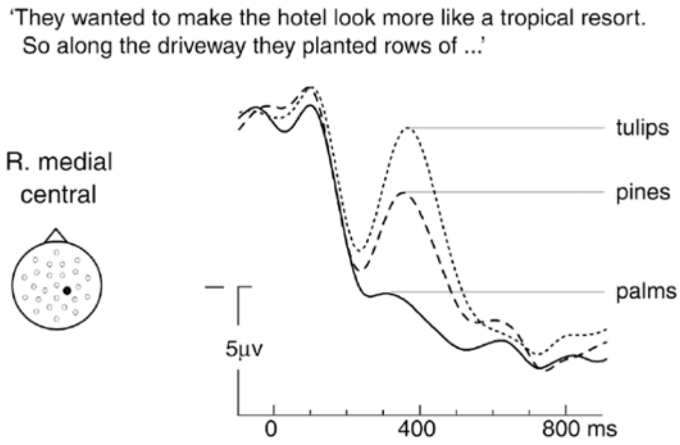

Figure 7. The influence of semantic memory structure on the N400. Averaged ERP waveforms at a representative electrode (CP2) exhibiting the smallest N400 amplitude for the expected continuation ("palms") and a greater reduction for words that pertain to the same semantic category ('pines') than for those that are semantically associated but belong to a different category ('tulips'). Two different mechanisms could explain this effect: lexical facilitation of within-category words ("pines") via the pre-activation of shared features (prediction view) or, alternatively, harder integration of between-category words ("tulips") due to poorer fit between the context and the encountered word (integration view). Negative is plotted upward. *Adapted from Federmeier & Kutas (1999).*

3.1.3.1 Functional interpretation of the N400 context effect

A piece of information that can contribute in understanding whether the N400 context effect can be taken as a sign of prediction is to determine where it is positioned in the cascade of processes triggered after word perception. As previously explained, there are two hypothesized ordered stages in language processing, (i) accessing the meaning of the word and (ii) integrating it with the sentence context. To some, the N400 context effect arises at a lexical stage when the meaning of the word is accessed. On such a view, the N400 context effect is due to predictable words being easier to retrieve from semantic memory (lexical view) (Federmeier & Kutas, 1999; Marta Kutas & Federmeier, 2000). For others, the N400 context effect results from a post-lexical stage (i.e. after the word has been accessed) when integration or "semantic unification" occurs (post-lexical view), and it is due to unpredictable words being more difficult to integrate (Brown & Hagoort, 1993; Hagoort, Baggio, & Willems, 2009)

3.1.3.2 Brain regions associated with the N400 context effect

One strategy to tease apart the cognitive processes underlying the N400 context effect is to pinpoint the brain networks that are likely to be involved in its origination (**Box 2**) (reviewed in

Lau, Phillips, & Poeppel, 2008; van Petten & Luka, 2006). fMRI studies using semantic priming paradigms that lead to N400 context effects have consistently found effects in the posterior medial temporal gyrus (pMTG) (**Figure 8A**). Superior and middle temporal areas have been linked with long-term storage of lexico-semantic representations (Binder, Desai, Graves, & Conant, 2009) as, for example, lesions in this area hinder performance in tasks that require access to lexical representations (Hart & Gordon, 1990). A consistent observation is that these regions show reduced activation for words with semantically related primes (e.g. tropical – palms) compared to those with unrelated primes (e.g. tropical – pines) (Lau, Gramfort, Hämäläinen, & Kuperberg, 2013; Lau & Namyst, 2019), which is consistent with reduced N400 amplitudes for primed words. In convergence with this, the N400m effect (Halgren et al., 2002) – the magnetoencephalographic (MEG) counterpart of the N400 context effect – has also been source-localized within the left mid posterior MTG (and other neighboring regions) (Service, Helenius, Maury, & Salmelin, 2007).

In turn, the involvement of the pMTL is less robust when sentence contexts are employed (instead of single-word contexts). Another region that is implicated when using sentence contexts is the left inferior frontal cortex (LIFC) (**Figure 8B**). The LIFC shows superior activity for semantic and/or pragmatic anomalies (Baumgaertner, Weiller, & Büchel, 2002; Kiehl, Laurens, & Liddle, 2002). Likewise, in MEG studies, the LIFC shows increased activation for incongruent than congruent words (Halgren et al., 2002; Pylkkänen, Oliveri, & Smart, 2009). Several authors have argued that the LIFC is concerned with semantic control, such as "controlled retrieval" (Wagner, Paré-Blagoev, Clark, & Poldrack, 2001) and "controlled selection" between competing representations (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). Compatible with this view, Hagoort, Baggio, & Willems (2009) attributed the activation of the LIFC to the recruitment of a "semantic unification network" tied to the integration of words into the context. Thus, post-lexical processes are also involved in the N400 context effect, especially in the face of semantically incongruent words or when there are multiple competing representations.

Box 2. A functional anatomic model of semantic processing in language

Lau and colleagues (2008) proposed a functional anatomic model of semantic processing in language grounded in psycholinguistic theory to help elucidate the functions most likely associated with N400 context effects. According to the model, the middle temporal gyrus (MTG) matches the phonemic or orthographic input with the stored lexico-semantic representation. Then, the anterior temporal cortex (ATC) and angular gyrus (AG) orchestrate the integration of new input with the ongoing contextual representation. The anterior inferior frontal gyrus (aIFG) and the posterior inferior frontal gyrus (IFG) utilize top-down information to mediate controlled retrieval and the selection of representations, respectively.

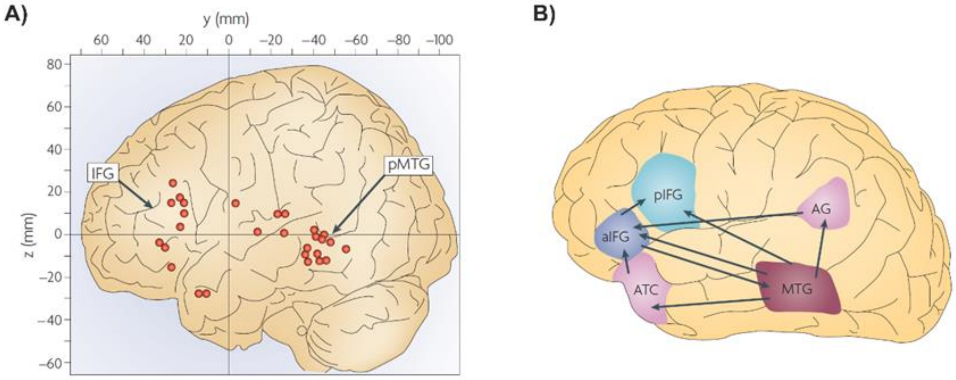

Figure 8. Pictorial representation of the main brain regions implicated in semantic processing of words in context and their hypothesized functional connections. A) Idealized brain depicting the approximate locations of the centers of significant activation in semantic priming studies (reviewed in Lau, Phillips, & Poeppel, 2008). Most effects gather around the left inferior frontal gyrus (IFG) and the posterior medial temporal gyrus (pMTG). **B)** Schematic representation of a functional neuroanatomic model for semantic processing proposed by the same authors. *Adapted from Lau, Phillips & Poeppel (2008).*

3.1.3.3 Multi-generator accounts of the N400 context effect

More recent studies have directly attempted to untangle the contribution of prediction from that of integration on N400 context effects. For example, a recent study found overlapping and independent effects of predictability and plausibility on the N400 (Nieuwland et al., 2019). Predictability is assumed to be a measure of prediction, and plausibility a measure of later integration demands (plausible words are easier to integrate than implausible words). Importantly, however, each effect had a different spatio-temporal profile – while predictability dominated the earlier part of the N400 context effect (starting as early as 200 ms post-word-onset), the effect of plausibility began later (about 350 ms post-word onset) and continued until about 650 ms after word onset. Similarly, another study reported that lexical prediction effects preceded by about 100 ms the effect of other forms of contextual facilitation (i.e. semantic overlap and plausibility) (Brothers, Swaab, & Traxler, 2015).

Interestingly, the seemingly later influence of integration may explain why N400 congruity effects sometimes extend substantially in time, lasting up to 800 ms in some cases: it may reflect sustained efforts to integrate an incongruent word with the prior context. Although non-significant, Lau, Namyst, Fogel, & Delgado (2016) also noted a tendency for congruity effects to begin later than those of predictability. In addition, this study found that when both factors were manipulated independently, the effect of predictability on the N400 amplitude was larger, more reliable, and had a different scalp distribution than that of congruity. According to the authors, this pattern of results is compatible with a multiple generator account of N400 context effects, whereby predictability effects capture facilitated lexical and semantic access and congruity effects result from a distinct process. This 'hybrid' view of the N400 context effect has increasingly gained theoretical and empirical support in the last years (Baggio & Hagoort, 2011; Kutas & Federmeier, 2011; Lau et al., 2016; Nieuwland et al., 2019) and it is also attractive from a neurophysiological standpoint – given the relatively long duration of the N400 context effect (from about 200 ms up to even 700 ms after word onset), there is enough time for multiple neural events to be imprinted in the EEG signal.

3.1.3.4 Strong forms of prediction

In order to circumvent the ambiguity between prediction and integration that comes with studying N400 context effects elicited by the critical word, other studies have sought evidence of pre-activation before the critical word is presented.

In the ERP realm, Wicha and colleagues (Wicha, Bates, Moreno, & Kutas, 2003; Wicha, Moreno, & Kutas, 2004) pioneered a strategy that consists of examining ERP differences at the word that precedes the critical word (pre-target word). For example, in Spanish, articles are gender-marked (feminine or masculine), such that it is possible to manipulate the gender-agreement between the article and the following noun (un/una canasta). Capitalizing on this, Wicha and colleagues (2003) showed that articles that disagreed with the expected noun (un canasta) yielded a larger N400-like response than those that matched the gender (una canasta). This suggested that, by the time participants encountered the article, they had already predicted gender features of the expected noun (for similar findings in Dutch, see Otten & Van Berkum, 2008; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). Besides morphosyntactic features, a study by DeLong and colleagued (2005) suggested that specific phonological representations can become pre-activated as well. In English, a different indefinite article is chosen based on whether the following noun starts with a consonant ('a kite') or with a vowel ('an apple'). Articles that mismatched ('an kite') elicited a larger negativity than for those that matched the expected noun ('a kite'). However, a recent large-scale study failed to replicate this finding (Nieuwland et al., 2018; see also Nicenboim, Vasishth, & Rösler, 2020). There is similarly mixed evidence regarding the pre-activation of specific orthographic forms, while some have found supporting evidence (Laszlo & Federmeier, 2009), others have argued that word form-information might be pre-activated only when there is sufficient time to predict the word (about 700 ms or more) (Freunberger & Roehm, 2016; Ito, Corley, Pickering, Martin, & Nieuwland, 2016; but see DeLong, Chan, & Kutas, 2020).

Taken together, there is evidence that readers and listeners predict at least down to the level of semantic and morpho-syntactic features (for more recent approaches showing convergent results, see Wang, Kuperberg, & Jensen, 2018). On the other hand, evidence for the pre-activation of phonemes or orthographic forms is less robust and might occur only in certain circumstances. Yet, a critical limitation of these effects is that, like N400 effects at target words, they capture the result of the prediction. That is, they do not capture the moment when predictions first arise and representations are first pre-activated, a point to which we will return later.

3.1.4 The costs of (mis)predicting

The empirical landscape that has been presented so far rebuts the initial reluctance with which prediction was met in psycholinguistics. Many linguists traditionally denied a role for prediction in language arguing that natural language is too unconstrained for predictions to be accurate. As Jackendoff (2002) put it, "*One might well predict that what comes after "little" in "the big star's beside a little…" is likely to be a noun […] but that still leaves open some tens of thousands of choices*". With more chances of being wrong than right, predicting would be an unnecessary waste of computational resources. It would be better to stick to an incremental strategy instead, processing input as it flows in, avoiding the risk of unsuccessfully getting ahead of it. As we have seen, there is now evidence that language is predictable and that even specific low-level aspects of words can be pre-activated. Notwithstanding this, some researchers have argued that unlike in other domains of cognition in which predicting is considered equivalent to processing (Clark, 2013), prediction may still not be a core operation in language processing (Huettig, 2015; Huettig & Mani, 2016; Pickering & Gambi, 2018). While prediction provides a 'helping hand' in some situations, it would not be necessary for language processing (Huettig, 2015; Huettig & Mani, 2016).

This caution is in part motivated by evidence suggesting that at least some forms of language prediction may be strategic, rather than automatic in nature. A long-standing notion in cognitive science is that there are two routes of processing: one that is automatic and implicit, and another that is controlled, more effortful, and typically demands attention (Posner & Petersen, 1990). In line with this, two mechanisms have been postulated to account for pre-activation in language comprehension: through fast and passive 'spreading activation' across structural links between pre-stored representations (Collins & Loftus, 1975) (i.e. priming) or through top-down mediated pre-activation based on the high-level context representation, which is what we understand by prediction in this thesis. Traditionally, priming has been deemed automatic and non-strategic, whereas prediction has been considered controlled and strategic (reviewed in Huettig, 2015; Kuperberg & Jaeger, 2016). In line with prediction being subject to strategic processes, one study investigated the influence of task instructions on the elicitation and the magnitude of N400 context effects (Brothers, Swaab, & Traxler, 2017). In one condition, participants were asked to simply read for comprehension, and, in another condition, they were told to predict the final word of the sentence. The N400 cloze-probability effect was larger in the latter condition, suggesting that prediction is mediated by global, top-down influences, such as task goals. In fact, they showed differences in the generation of lexical prediction cannot be explained by word-level factors (co-occurrence or lexical association), but that top-down mechanisms intervene in dialing up or down the degree to which lexical and semantic pre-activation takes place. Another piece of evidence that points to a strategic use of prediction comes from the finding that effects of predictability are reduced or eliminated when the predictive validity of the task is low, that is, when predictive cues are invalid (Brothers et al., 2017), or when the broader context encourages prediction to a lesser extent (Lau, Holcomb, & Kuperberg, 2013).

Strategic processing is typically considered more resource-demanding because it requires controlled guidance and monitoring. Consequently, a strategic use of prediction suggests that generating predictions may be costly, and therefore that it may not be always optimal (or possible) to compute them. In line with this, individual differences in cognitive abilities may be a mediating factor of predictive language processing as well. For instance, several studies have demonstrated that age-related decline reduces the facilitation of contextually predictable words (Federmeier, 2007; Federmeier & Kutas, 2005; Federmeier, McLennan, Ochoa, & Kutas, 2002; Wlotko, Federmeier, & Kutas, 2012). Older adults seem to be less apt at taking advantage of semantically rich contexts to facilitate subsequent processing (Federmeier & Kutas, 2005). Also, ERP patterns consistent with prediction are significantly reduced in older adults with lower verbal fluency scores (associated with reduced cognitive resources) (**Figure 9**) (Federmeier et al., 2002; see also Federmeier, Kutas, & Schul, 2010). Based on these findings, Federmeier and colleagues speculated that predictive language processing may be too cognitively-taxing for older adults (especially those with fewer cognitive resources), who would prefer a 'laissez faire', incremental strategy instead (reviewed in Federmeier, 2007). Beyond age, other individual factors that are linked with general cognitive ability have been suggested to affect language prediction, including working memory (WM) capacity or processing speed (for a review, see Huettig, 2015).

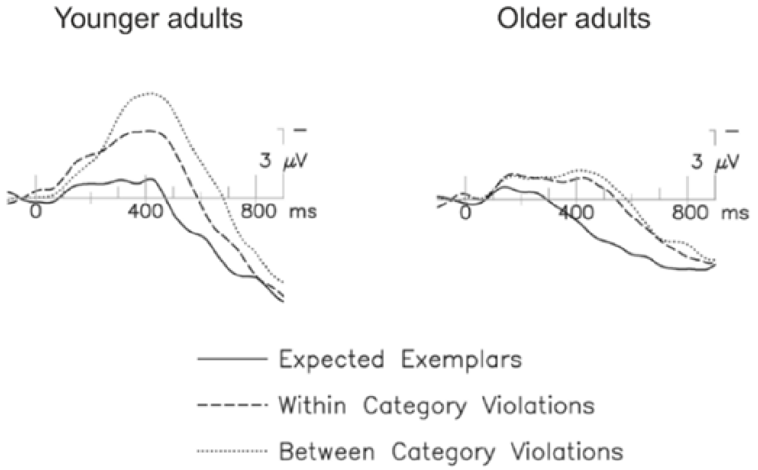

Figure 9. Age-related changes in averaged ERP patterns consistent with prediction. Younger adults show smaller N400 amplitudes for within- than for between-category candidates, whereas older adults show statistically undistinguishable responses. Under a prediction view, this is consistent with older adults being unable or less prone to pre-activating expected features. Negative is plotted upward. *Adapted from Federmeier & Kutas (2005).*

Besides potential processing costs associated with computing the predictions themselves, another type of cost that would be intrinsic to predictive processing is that of handling failed predictions – in the end, prediction is, by definition, fallible. In this respect, several ERP studies have observed that, in addition to larger N400 responses, words that violate contextual predictions also elicit later positive-going ERP component about 500 to 1000 ms post-word onset (Delong, Urbach, Groppe, & Kutas, 2011; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007b; Thornhill & Petten, 2012). The timing and the scalp distribution of these late positivities varies across studies, but a pattern that seems reliable is that unexpected but congruent words elicit frontal positivities (e.g. "Don't touch the wet dog"), whereas unexpected and anomalous words produce posteriorly distributed positivities (e.g. "Don't touch the wet star") (**Figure 10**) (Delong, Troyer, & Kutas, 2014; Kuperberg et al., 2019; for a review, see van Petten & Luka, 2012). These positive-going waves have been linked to different computations triggered in the face of misprediction, including the suppression of an unfulfilled prediction (Federmeier et al., 2007a), the detection of conflict between the predicted and the actual input (Delong et al., 2011), or an effort to rectify the interpretation (Brothers et al., 2015). Finally, under predictive processing frameworks (see page 20), late positivities have been hypothesized to reflect an update of the situation model (Kuperberg et al., 2019) or conflict detection and subsequent revision/repair the interpretation (Kuperberg et al., 2019; Rabovsky et al., 2018).

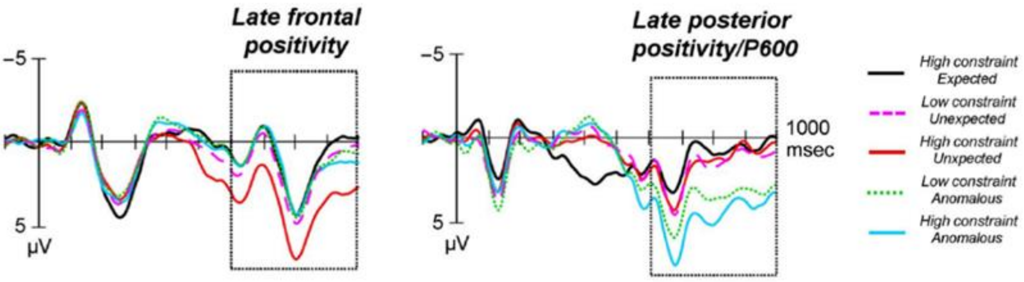

Figure 10. Averaged ERPs exhibiting the late positivities in response to the violation of contextual predictions in sentence comprehension. On the left, averaged ERP at a frontal location (FPz electrode) exhibiting the late frontal positivity elicited by unexpected words in strongly predictive contexts (red line). On the right, at a posterior location (Oz electrode), the late posterior positivity evoked by semantically anomalous words regardless of contextual constraint. *Adapted from Kuperberg, Brothers, & Wlotko (2019).*

3.1.5 Interim summary

Until here, we have reviewed key pieces of evidence in support of prediction in language comprehension. As a sentence unravels, readers and listeners rapidly incorporate new words into the preceding context to construct meaning. Word predictability matters in this operation, such that words that are more predictable are easier to process, as evidenced by faster reading times and reduced amplitudes of the N400 response (i.e. N400 context effect). This facilitation is taken to reflect prediction, whereby predictable words are easier to process thanks to pre-activation. However, it can also be explained due to predictable words being easier to integrate with the current sentence representation. Accruing evidence points to the contribution of both prediction and integration processes to word predictability effects, albeit prediction seems to have an earlier impact. In addition, in some cases, words that violate contextual predictions elicit late positivities that are proposed to capture the consequences of misprediction. Overall, several studies support that different types of information, from semantic down to morphosyntactic features, can be predicted during sentence comprehension. Embracing prediction as a fundamental operation in language, computational models for predictive language processing have been postulated. Yet, there is also evidence pointing to important limitations as to what is predicted and when, as well as to who engages, or can engage, in predictive language processing.

3.2 Before finding out: the anticipatory stage

As reviewed in previous sections, the idea that certain aspects of information are pre-activated, at least under some circumstances, is now well-established. However, most available evidence to date has focused on neural activity elicited by critical linguistic input (either predicted words themselves or anteceding words). To the extent that information gets pre-activated in advance, there should be correlates of brain activity associated with prediction already before the input is perceived. Following this rationale, the current research aimed to explore electrophysiological signatures that could capture the anticipatory stage of prediction, that is, the stage before linguistic predictions are tested against bottom-up input.

The division between an anticipatory and a processing stage in language prediction is neatly illustrated by predictive processing frameworks (as depicted in **Figure 5**, page 16). Accordingly, most existing ERP evidence in the domain of sentence comprehension, which has focused on post-stimulus activity, would correspond to the processing or learning stage. In turn, little is known about the mechanisms that support language prediction in the anticipatory stage would encompass neural activity associated with the generation of such predictions. While this work remains agnostic about how language prediction is specifically implemented by the brain, it is clear that our understanding of the mechanisms that support predictive language processing will be incomplete until reliable markers of prediction in this anticipatory stage are available.

Following this reasoning, the current work sought to explore electrophysiological indices of predictive language processing *before* critical words are perceived. From here onwards, we will use the term anticipation to broadly refer to the cognitive operations engaged in preparation for critical words, and that result from the generation of predictions. As a starting point in this quest for a neural correlate that captures such operations, we focused on slow cortical potentials (SCPs), a family of electrophysiological signatures that have been consistently linked to prediction and anticipation in other cognitive domains. In the remainder of the Introduction, I will introduce the general features of SCPs and the most well-known ERP components of this kind. Finally, I will briefly turn to the domain of neural oscillations, another type of electrophysiological phenomena that can provide additional and complementary information to that contained in ERPs.

3.2.1 Slow negative potentials as indices of anticipation

In the domain of ERPs, a family of components tightly linked to anticipation in a variety of domains are slow cortical potentials (SCPs). SCPs are long-lasting deflections in the EEG signal, extending from hundreds of milliseconds up to several seconds. Most of these SCPs tend to have a negative polarity relative to stimulus baseline and to emerge gradually in the fore period of certain events, such as the execution of a movement (Kornhuber & Deecke, 1965; Walter, Cooper, Aldridge, McCallum, & Winter, 1964), or the presentation of a motivationally salient stimulus (C. H.M. Brunia & Damen, 1988).

Slow negative potentials (SNP) that have a duration of several hundreds of milliseconds and broad scalp distributions (4-5 cm2) originate most probably from synchronized excitatory postsynaptic potentials within cortical structures underneath their recording site (for an extensive discussion, see Birbaumer, Elbert, Canavan, & Rockstroh, 1990; Brunia, van Boxtel, & Böcker, 2012). Hence, the elicitation of a SNPs at a particular electrode site is taken as a manifestation of an enhanced functional state in the underlying cortical area, relative to other areas or to other temporal intervals. SNPs typically surface over task-relevant regions (Khader, Schicke, Röder, & Rösler, 2008; Rösler, Heil, & Röder, 1997), with an amplitude that is modulated by a variety of cognitive factors. Next, I will briefly review distinct SNPs that have been described in the context of anticipating relevant stimuli and events, highlighting their most important features.

3.2.1.1 The Contingent Negative Variation (CNV)

In 1964, Grey Walter and colleagues published one of the first demonstrations of a systematic SCP in humans (Walter et al., 1964). They presented participants with a “warning stimulus” (S1; a single click) that announced the presentation of an “imperative” stimulus (S2; repetitive flashes), to which a motor response was required. They found that, before the presentation of S2, a distinct electrophysiological response arose – the *Contingent Negative Variation* (CNV) (**Figure 11A**), an increasing negative SCP at frontocentral sites that reached its peak before S2, and then returned to baseline after the motor response was made. Intriguingly, the CNV was not present when no motor response was required, strongly suggesting that it reflected motor operations. Yet, it did not simply depend on movement execution – a CNV was also elicited when participants were prompted to perform a mental task upon S2, with no overt response. Thus, although motor aspects were a sufficient condition to elicit a CNV, they were not a necessary condition. What was mandatory was that S1 and S2 were somehow associated with each other. If S1 and S2 were presented in isolation, or successively but with no contingent relation, then the CNV was absent. Based on its features, the CNV was taken as a sign of "sensory-motor association and expectancy" (Walter et al., 1964) – indeed, it is sometimes referred to as the "expectancy wave".

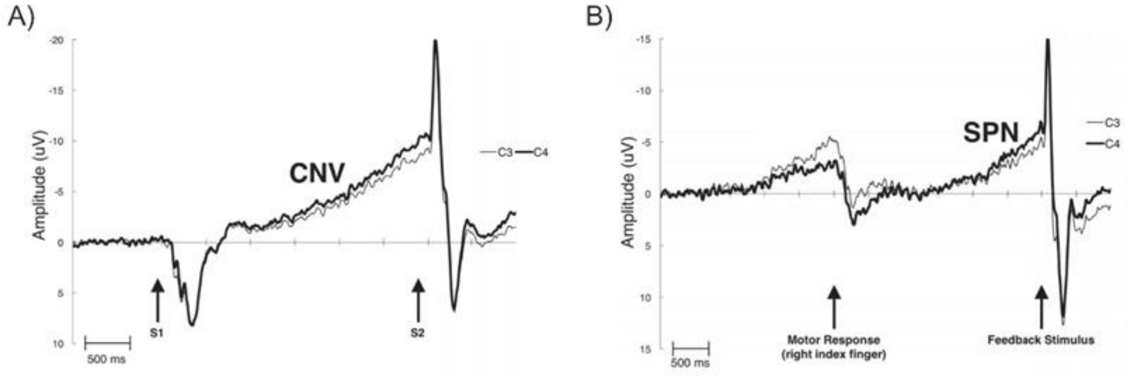

Figure 11. Contingent negative variation (CNV) and Stimulus-preceding negativity (SPN). A) Averaged ERP waveform showing the CNV in the interval between the warning stimulus (S1) and the imperative stimulus (S2), larger at right (C4) than left (C3) electrodes. **B)** Averaged ERP waveform showing the gradual development of the SPN prior to the presentation of the feedback stimulus, also larger at right (C4) than left (C3) sites. Negative is plotted upward. *Adapted from Kotani, Ohgami, Yoshida, Niryu & Inoue (2017).*

Over the years, a bulk of research has sculpted a detailed picture of the CNV. A major contribution was made by a series of studies inserting longer intervals between S1 and S2. This simple manipulation revealed that the CNV can be subdivided into two distinct phases (Connor and Lang, 1969). An early phase (CNVe), composed of a frontal negative peak, and a terminal phase (CNVt) consisting of a centrally distributed negativity that peaks right before S2. The CNVe was seen as an orienting response to S1 ("orienting wave" or "O-wave"), whereas the later CNVt was related to the anticipation of S1 ("expectancy wave", or "E-wave") (Loveless & Sanford, 1974).

A few years later, Rohrbaugh & Gaillard (1983) advocated for a purely motor interpretation of the CNVt. This was greatly motivated by the strong resemblance between the CNVt and the Readiness Potential (RP). The RP (originally coined in German as Bereischaftpotential, BP; Kornhuber & Deecke, 1965) is a negative SCP that precedes the execution of a voluntary movement, instead of an externally cued movement as in the case of the CNVt. Their similarities in terms of morphology and scalp distribution opened a reasonable debate on whether the two components were indeed the same, with convincing arguments both for and against (for a detailed enumeration, Brunia &

Van Boxtel, 2004). However, as earlier studies had already suggested (Walter et al., 1964), the CNVt is also observed even in the absence of a motor response (D. S. Ruchkin, Sutton, Mahaffey, & Glaser, 1986). Indeed, as we will see next, motor and nonmotor aspects of the CNV can be teased apart.

3.2.1.2 The Stimulus Preceding Negativity

In the 80s, several studies alluded to the existence of nonmotor negativities preceding relevant stimuli. In an experiment by Grünewald-Zuberbier, Grünewald, Runge, Netz, & Hömberg (1981), participants had to perform skilled movements and then received visual feedback about the correctness of their movement. Prior to the presentation of the feedback, a right-lateralized negative SCP arose, independent of the hand used to respond. A few years later, Damen and Brunia (1985; 1987) replicated this finding and coined the component *Stimulus-Preceding Negativity* (SPN) (**Figure 11B**). In particular, they used a time-estimation paradigm, in which S1 prompts participants to press a button after a given time (e.g. 3 seconds) that they need to estimate mentally. Then, a fixed time after their response, they receive feedback (S2) about the accuracy of their estimation (i.e. "correct", "too early", or "too late"). The SPN is reliably observed prior to the feedback stimulus, in the absence of a motor response (for a review, see Brunia, Hackley, van Boxtel, Kotani, & Ohgami, 2011). For the SPN to emerge, a necessary condition is that (i) attention is directed to the stimulus, and that (ii) the stimulus provides useful information. The motivational aspect of the anticipated stimulus plays a critical role as well. When S2 simply provides instructions about the task at hand, the SPN is smaller and not as reliable compared to when S2 delivers performance-related feedback (Damen & Brunia, 1994; Kotani & Aihara, 1999).

The SPN has been interpreted as reflecting the level of expectancy for the outcome (Brunia et al., 2011). This fits well with the finding of a larger SPN prior to a stimulus that will determine if there is a monetary gain, compared to when the stimulus is not decisive in this respect (Donkers & van Boxtel, 2005). However, in other cases, the SPN follows a pattern that is harder to accommodate with this interpretation. Several studies have found that the SPN is larger for probabilistically unpredictable outcomes than for probabilistically predictable ones (**Figure 12A**) (Catena et al., 2012; Foti & Hajcak, 2012; Fuentemilla et al., 2013). An alternative functional interpretation that can parsimoniously account for these apparently conflicting results is that the pre-freedback SPN reflects the amount of information that one expects to extract from the incoming stimulus. For instance, Morís, Luque, & Rodríguez-Fornells (2013) tracked learning-related changes in the SPN during a reinforcement learning task, in which participants learn the relationship between stimuli by means of making guesses and receiving feedback accordingly. They showed that, as learning progressed, the amplitude of the SPN became smaller (**Figure 12B**) (Ren, Valle-Inclán, Tukaiev, &

Hackley, 2017). If the SPN reflected outcome expectancy (Brunia et al., 2011), then its amplitude should increase as the outcome becomes more predictable with learning. In turn, the observed learning-induced decrement may be due to the gradual loss of informational value of the feedback as the associations are internalized. In line with this hypothesis, Kotani and colleagues (2003) found that, in trials with highly informative feedbacks, the SPN showed larger amplitude than when the feedbacks were uninformative.

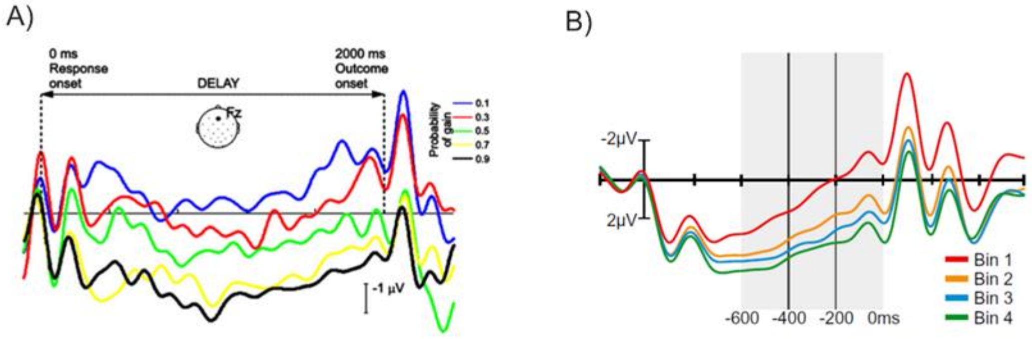

Figure 12. Amplitude modulations of the pre-feedback SPN. A) Averaged ERP at the Fz electrode exhibiting the SPN during the anticipation of feedback informing about whether money was gained, or not. The task was divided in blocks, and the probability of gaining money in each block varied from 0.1 to 0.9 (each colour corresponds to the assigned probability). The amplitude of the SPN was linearly related to the probability of gain, such that the SPN was larger for probabilistically unpredictable than predictable outcomes. Data was low-pass filtered at 6 Hz. *Adapted from Fuentemilla et al. (2013).* **B)** Learning-induced modulations of the SPN amplitude during reinforcement learning. Averaged ERPs at the Fz electrode showing the progressive reduction of the pre-feedback SPN (from −600 to 0 ms time-locked to feedback onset) as learning progressed. To capture the effect of learning, each bin (Bins 1 to 4) was composed of a quartile of the training trials in consecutive learning blocks (e.g. Bin 1 contained the first 18 trials of the first block, and so forth). *Adapted from Morís, Luque, & Rodríguez-Fornells (2013).*

Another way to explain these apparently dissimilar findings is to understand the SPN as the common electrophysiological signature of a family of different waves tied to anticipation. Besides feedback stimuli, there are other types of stimuli that are preceded by an SPN (for a review, see van Boxtel & Böcker, 2004). As mentioned earlier, although smaller, stimuli that provide instructions about the task at hand also induce an SPN (van Boxtel & Brunia, 1994). A factor that strongly modulates the observation of an SPN is the affective valence and salience of the imperative stimulus. Naturally, the anticipation of a stimulus that carries information about consequences for oneself must have an affective quality and, accordingly, the SPN is found prior to many different affective stimuli of both positive and negative valence (Böcker, Baas, Kenemans, & Verbaten, 2001; Kotani, Hiraku, Suda, & Aihara, 2001; Poli, Sarlo, Bortoletto, Buodo, & Palomba, 2007). Interestingly, the SPN elicited by these stimuli shows slightly different scalp distributions. Albeit still right-hemisphere preponderant, the SPN is maximal over distinct areas along the anterior-to-posterior plane, suggesting the involvement of different frontal generators. In addition, other studies have found sensory modality-specific scalp distributions. For example, a larger negative SCP is found over occipital sites prior to visual feedback than prior to auditory feedback (Brunia & Van Boxtel, 2004). Further evidence corroborates that the cortical distribution of the SPN is not unitary, but it is influenced by a variety of parameters, including the experimental task, the anticipated content, or the psychological state of the participant (Kotani et al., 2015)Altogether, rather than a unitary phenomenon, the SPN is most likely the physiological expression of a variety of processes that have their anticipatory character as a common denominator.

3.2.1.3 A potential link between slow negative potentials and prediction in language

Interestingly, there a few instances that hint to the emergence of SNPs in relation to the anticipation of words during sentence comprehension. A common procedure to examine brain responses during the comprehension of written sentences is the rapid visual serial presentation (RVSP), wherein sentences are presented one word at a time on a screen, allowing to track how processing evolves word-by-word. Using this method, van Petten & Kutas, (1991) contrasted the ERP response of content and function words in congruent sentences, syntactically anomalous sentences, or random word strings. In congruent sentences and syntactically anomalous sentences (but not in random word strings), function words elicited a larger SNP than content words about 400 to 700 ms post-word onset (which they referred to as N400-700) (**Figure 13A**), which was cataloged as a type of CNV – they argued that function words, the class of words that elicited the negativity, might have cued the forthcoming presentation of a more informative, content word.

Also pursuing a different goal, Besson, Faita, Czternasty, & Kutas (1997) employed a particularly interesting experimental design to assess the anticipatory stage of prediction. In fact, the studies in the current dissertation follow a similar logic. They manipulated the expectancy of sentence-final words through contextual constraint (proverbs or low constraining contexts) and introduced a 600 ms pause before the sentence-final word in half of the trials, thus emulating S1-S2 paradigms typically used to study anticipation. They found that, during the pause, a negative SCP developed for low constraining contexts (compared to proverbs), maximal at posterior and temporal sites (**Figure 13B**).

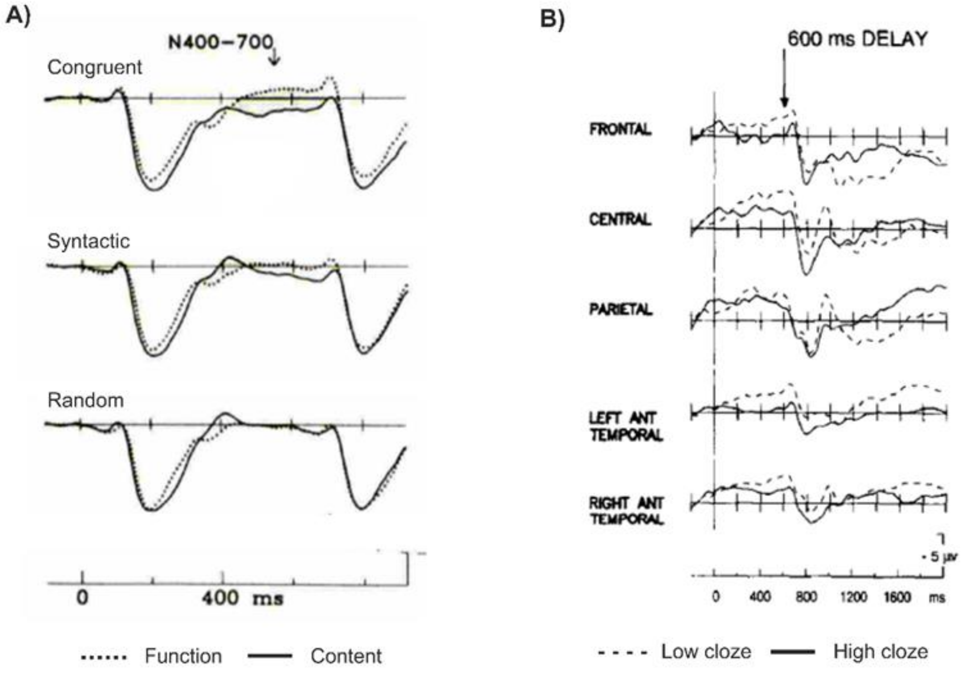

Figure 13. Previous instances of negative SCPs appearing between words during sentence processing. A) Word-locked averaged ERPs (±2 microvolts) at Fz electrode for different types of sentences. A negative SCP (N400-700 ms) develops after function words (relative to content words) in congruent sentences, it is present but smaller in syntactically anomalous sentences and absent in random word strings. Adapted from van Petten (1995). **B)** Averaged ERP waveforms for high and low cloze sentence-final words at different locations. A 600 ms delay was inserted between the penultimate and the sentence-final word. The arrow indicates the onset of the sentence-final word. During the delay, a larger negative SCP unfold for low than for high cloze endings. *Adapted from Besson, Faita, Czternasty & Kutas, (1997).*

Despite the differences in their experimental designs, these results are interesting because the amplitude of the SNP is seemingly modulated as a function of the level of expectancy that the sentence contexts afford about an upcoming word. Unlike associative learning tasks, in this case, the expectancy is not derived from arbitrary associations learned during the experiment, but it is based on probabilistic information extracted from the sentence context. Therefore, this shows potential for SNPs to be able to map activity associated with linguistic predictions during sentence comprehension.

3.2.1.4 Slow negative potentials in sentence comprehension

More broadly, sustained SNPs are not an uncommon phenomenon in the sentence comprehension literature. In this realm, SNPs have been generally associated with demanding processing, as they are usually found when processing sentences with syntactically complex or ambiguous structures.

A case in point is the Nref, a sustained anterior negativity that is robustly observed in sentences with referential ambiguity, in which an anaphor ('the book') has more than one referent (i.e., there is more than one book) (Van Berkum, Brown, & Hagoort, 1999). Other situations that tend to elicit sustained late anterior negativities are the processing of filler-gap dependencies (Fiebach, Schlesewsky, & Friederici, 2001, 2002; J. W. King & Kutas, 1995; Kluender & Kutas, 1993; Matzke, Mai, Nager, Rüsseler, & Münte, 2002; Piai, Meyer, Schreuder, & Bastiaansen, 2013) or semantic violations regarding temporal relationships (Münte, Schiltz, & Kutas, 1998a; Politzer-Ahles, Xiang, & Almeida, 2017).

A common denominator across the situations that elicit sustained SNPs is that they are thought to be taxing for working memory (WM). The concept of WM is used to refer to the ability to maintain information available within the cognitive system and simultaneously manipulate this information (Atkinson & Shiffrin, 1971; Baddeley & Hitch, 1974). The notion that WM is involved in language comprehension is mainly rooted in the fact that language unfolds sequentially, so any cognitive operation beyond the processing of the current bottom-up input (e.g. going back to earlier parts of the sentence to re-elaborate the interpretation) requires a system that allows to maintain and manipulate information internally.

For example, sentences with object relative clauses, such as "the reporter who the senator harshly attacked admitted the error" (the relative clause is underlined) are usually difficult for readers (King & Just, 1991). This is theorized to be due to the object (i.e. 'who') appearing untypically before the subject ('the senator'). King & Kutas (1995) examined the ERP pattern evoked while reading these sentences, compared to sentences with easier, subject relative structures, like "the reporter who harshly attacked the senator admitted the error". A sustained anterior negativity developed for object relative to subject sentences over the early region of the relative clause (i.e., "the senator") (**Figure 14**) (see also Müller, King, & Kutas, 1997), which was taken to reflect increased WM demands devoted to the temporary maintenance of linguistic elements that could not be immediately utilized. Specifically, theory posits that the *displaced* object, the filler, is temporarily stored in WM until the position where it should have appeared, the gap, is reached (i.e. after 'the senator'), the moment in which it is integrated with the sentence (e.g., *Active Filler Hypothesis*; Clifton & Frazier, 1989). Further reinforcing a link between these negativities and WM, their amplitude correlates with the WM capacity of readers (Vos, Gunter, Schriefers, & Friederici, 2001).

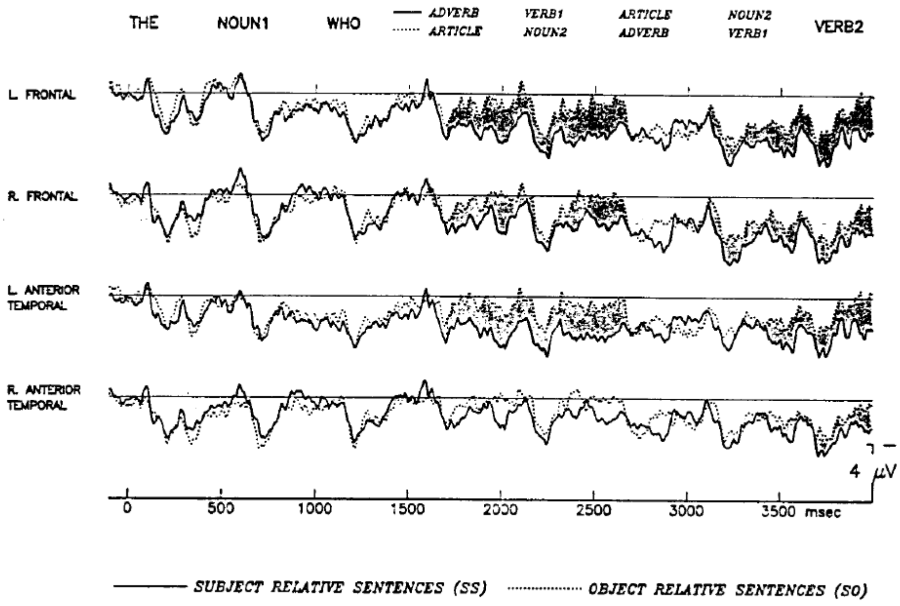

Figure 14. Sentence-level sustained anterior negativities associated with syntactic complexity. Grand-averaged ERPs for the multi-word waveform of subject and object relative sentences (i.e. canonical and non-canonical sentence structures). An anterior negativity developed for object relative to subject sentences over the early region of the relative clause (i.e., "the senator"), which was taken to reflect increased WM demands devoted to the temporary maintenance of linguistic elements that could not be immediately utilized. *Adapted from Kutas & King (1996).*

Interestingly, some of the most prominent theories to account for the processing of these syntactically complex structures assume, either implicitly or explicitly, the generation of predictions. Namely, the 'Syntactic Prediction Locality Theory' (Gibson, 1998), distinguishes between two components that determine the consumption of computational resources during comprehension: (i) the 'integration cost' component, and (ii) the 'memory cost' component. The integration cost component refers to the cognitive resources devoted in the moment of integrating new words into the sentence representation, whereas, interestingly, the "memory cost component" specifies cognitive resources involved in the temporary storage of activated linguistic input based on syntactic predictions. Importantly, processing costs are greater the longer the predicted input needs to be maintained in WM, and the longer the distance between the elements to be integrated. Fiebach and colleagues (2001) found empirical support for the functional dissociation of these two components comparing questions with longer or distances between the filler and its gap. They found a sustained left anterior negativity for object questions with longer filler-gap distances, but not for short object questions. They interpreted this slow negativity as reflecting the maintenance of the filler object in WM (memory cost component), whereas integration costs were associated by transient positivities elicited at specific-word locations (integration cost component).

In addition to syntactic complexity, conceptual and pragmatic aspects can also lead to the development of SNPs (Münte et al., 1998; see also, Politzer-Ahles et al., 2017). For instance, sentences that begin with “before” present events in reverse chronological order, such that what happened first is revealed in second place, whereas the same sentence starting with “after” presents events in sequential order, matching our real-world experience whereby present events precede future ones (‘*Before/After* the psychologist submitted the article, the journal changed its policy’). Interestingly, Münte and colleagues (1998) found that the word “before” elicited left anterior SNPs relative to “after” already within 300 ms, a difference that increased across the sentence. In addition, the magnitude of the effect correlated with the WM span of participants, altogether pointing to high-level and conceptual knowledge (in this case about temporal order) as mediators of sentence processing demands as well.

3.2.2 Alpha and beta neural oscillations associated with anticipation

Besides ERPs, another aspect of electrical brain activity that can be extracted from EEG (or MEG) signal are neural oscillations – increases and decreases in power (i.e. wave amplitude) that are thought to reflect, respectively, synchronization and desynchronization of firing patterns in neural populations (Buzsáki & Watson, 2012). The standard approach to evaluate oscillatory activity are time-frequency (TF) analyses, which allow quantifying power at multiple frequencies over time, including delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (>30 Hz) bands. Different frequencies have been tied distinct cognitive functions, including language processing (for reviews, see Bastiaansen, Mazaheri, & Jensen, 2012; Meyer, 2018; Piai & Zheng, 2019). Interestingly, unlike ERPs, neural oscillations capture activity that is not phase-locked to the time-locking event (Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996), and thus can reveal additional and complementary information about the same underlying process.

In other cognitive domains, alpha and beta oscillatory activity has been associated with prediction and anticipation (Arnal & Giraud, 2012). Alpha oscillations have been proposed to serve as a mechanism to gate relevant sensory information (Jensen & Mazaheri, 2010), whereby relevant information is prioritized. According to the gating-by-inhibition model (Jensen, Bonnefond, & Vanrullen, 2012), alpha power increases reflect the disengagement of task-irrelevant regions through inhibition to protect relevant internal representations from interfering or potentially interfering stimuli. By opposition, alpha power decreases indexes enhanced cortical engagement of task-relevant areas (Klimesch, 2012). For example, in visuospatial tasks, when participants attend to the left hemifield, alpha power decreases at posterior sites in the contralateral right hemisphere and it concomitantly increases in the ipsilateral hemisphere (**Figure 15**) (Händel, Haarmeier, & Jensen, 2011; Rihs, Michel, & Thut, 2007; Worden, Foxe, Wang, & Simpson, 2000). Further supporting a functional role of alpha oscillations in sensory processing, pre-stimulus alpha power decreases at occipital electrodes predict better visual discrimination of the imperative stimulus (Klimesch, Sauseng, & Hanslmayr, 2007; Mathewson et al., 2014; Mazaheri et al., 2014). The evidence of auditory alpha oscillations is scarcer, but there is also evidence of task-specific alpha suppression in the auditory domain linked to anticipation of the target stimulus (for a review, Weisz, Hartmann, Müller, Lorenz, & Obleser, 2011). For example, alpha is suppressed in the contralateral to-be-attended stream (over temporal sites) during a speech monitoring task (Kerlin, Shahin, & Miller, 2010), and these lateralization effects have been found during the anticipation of auditory targets (Ahveninen, Huang, Belliveau, Chang, & Hämäläinen, 2013; Banerjee, Snyder, Molholm, & Foxe, 2011).

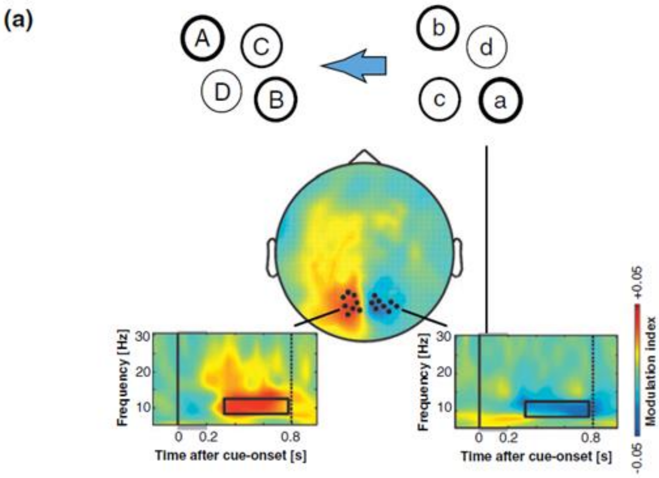

Figure 15. Hemispheric lateralization of alpha activity associated with attentional allocation. In a visuospatial attention task, a set of visual stimuli are presented concomitantly in the left and right hemifields (circled letters). When participants attend to the left hemifield (as the left-pointing arrow indicates), alpha power decreases at posterior sites in the contralateral right hemisphere and it simultaneously increases in the ipsilateral hemisphere. This activation pattern is taken to reflect attentional allocation, whereby alpha decreases index the engagement of task-relevant areas and alpha increases reflect functional inhibition. *Adapted from Jensen, Bonnefond & VanRullen (2012).*

On the other hand, alpha and beta power decreases have also been described in relation to memory operations, including encoding (Hanslmayr, Spitzer, & Bäuml, 2009) and retrieval (Khader & Rösler, 2011; Burgess & Gruzelier, 2000) of episodic memories, as well as WM retention (van Ede, Niklaus, & Nobre, 2017). For instance, in associative memory tasks, alpha power decreases over content-specific areas during retrieval, that is, at frontal electrodes when retrieving objects, and at parietal electrodes when retrieving spatial locations, even though in both cases the associations were learned and tested in the visual modality (Khader & Rösler, 2011; see also, Burgess & Gruzelier, 2000). Furthermore, alpha power decreases are modulated by the number of items to be retrieved (Khader & Rösler, 2011) or maintained in WM (van Ede et al., 2017; reviewed in van Ede, 2018). These results point to alpha power decreases being sensitive to the quality and quantity of the content that is being reactivated or retained.

Recent theoretical models view neural desynchronization at low-frequency bands, such as alpha and beta bands, as mediating the representation of information. Hanslmayr, Staudigl, & Fellner, (2012) simulated neural populations with different degrees of synchrony (**Figure 16**) (no synchrony, low synchrony, or high synchrony) and showed that a state of low synchrony is associated with the encoding of more specific information in the pattern of neural spiking, whereas, in turn, a state of high synchrony is related to less information being encoded. Low-frequency power decreases would mechanistically afford a more complex or information-rich state, which enables a neural population to express a stimulus-specific code. In support of this, alpha and beta power decreases track the fidelity of stimulus-specific information represented in the cortex (Griffiths et al., 2019). But rather than coding information itself, alpha and beta power decreases would confer windows of opportunity that boost information representation and information processing capacity. Note that this view is also compatible with the previously introduced proposal of alpha oscillations as reflecting functional inhibition – alpha power decreases might enhance information processing via functional dishinibition of relevant cortical networks.

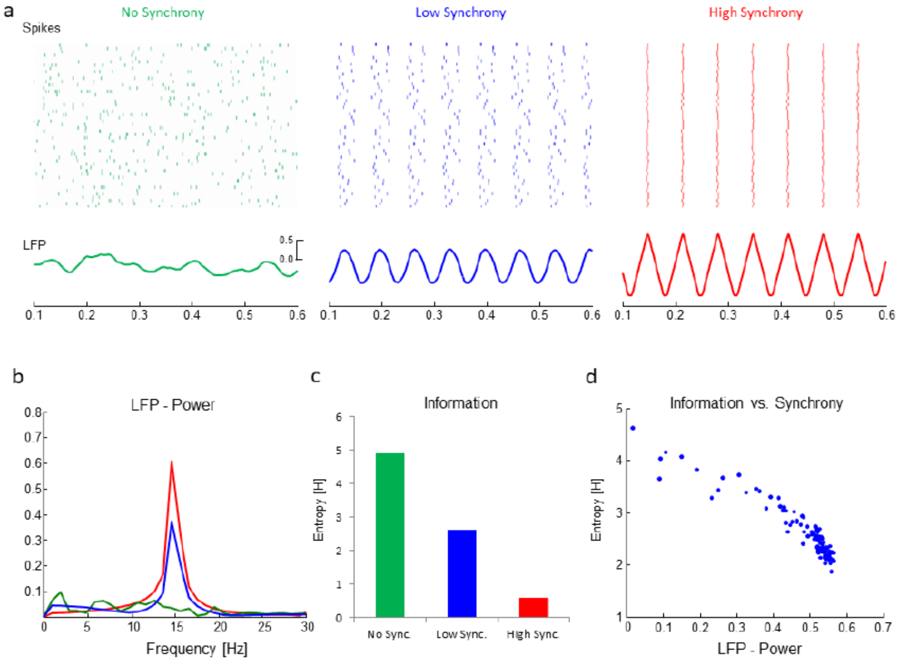

Figure 16. Information-via-desynchornization hypothesis. A) Simulated firing rates for a neural assembly with different degrees of synchrony (no synchrony, low synchrony or high synchrony, from left to right) with the same total number of spikes. The corresponding local field potentials are shown below each panel. B) As synchrony increases, the power of the local field potential also increases (at 15 Hz). C) Information obtained from the firing rates of each state of synchrony. D) Scatterplot showing the inverse relation between the power of the local field, varying in each simulation, and information. *Adapted from Hanlsmayr, Staudigl, & Fellner (2012).*

Returning to the domain of language, a paucity of studies has inspected anticipatory oscillatory activity in language. One of them was performed by Dikker & Pylkkänen (2013), who employed pictorial contexts. In their task, some pictures were predictive of one specific lexical candidate. For example, the picture of an apple would be predictive of the word ‘apple’, whereas others did not provide this univocal prediction, for example, the picture of a grocery bag did not allow to pin down a specific fruit. In predictive contexts, the final word was preceded by a theta power increase (4-8 Hz) that was associated with the top-down modulation of the visual cortex, suggesting that the expected word form was pre-activated. Although noteworthy, the extrapolation of these findings to sentence comprehension is limited due to the utilization of pictorial contexts. In sentence comprehension, some researchers have inspected anticipatory activity using the same tasks employed to study N400 context effects (see Section 3.1.2.1). One study found that sentence-final words in highly constraining contexts (e.g., “the children went out to *look*”) were preceded by an alpha power decrease over posterior-occipital sites, compared to low constraining ones ("Joyce was too frightened to *look*") (Rommers, Dickson, Norton, Wlotko, & Federmeier, 2017). In a similar fashion, another study reported pre-word alpha power decreases in high compared to low constraining contexts, but at frontal and posterior and left-lateralized in that case (Wang, Hagoort, & Jensen, 2018).

Finally, also of relevance regarding lexical and semantic pre-activation in comprehension, alpha and beta power (8-25 Hz) decreases are reliably found in association with lexical and semantic retrieval during speech production. In context-driven naming tasks, sentence contexts are followed by a depiction of the expected continuation (e.g. "She locked the door with the [picture of a key]") and participants must name the picture when it appears after a short interval. Prior to picture naming, alpha and beta power decreases in constraining relative to neutral sentences in temporal and inferior parietal (and sometimes frontal) brain areas (Piai, Meyer, Dronkers, & Knight, 2017; Piai, Roelofs, & Maris, 2014; Piai, Roelofs, Rommers, & Maris, 2015; Piai, Rommers, & Knight, 2018; for a review, see Piai & Zheng, 2019).

3.2.3 Interim summary

Most of the available evidence in support of prediction in language comprehension has focused on brain activity after the presentation of critical words. From the perspective of predictive processing, such post-stimulus activity likely represents the 'processing' or 'learning' stage, when the predicted content is compared with the actual input. In turn, only a few studies have focused on pre-stimulus activity in the context of language prediction. In the ERP domain, it is well-established that pre-stimulus SNPs are concerned with anticipation and memory operations, and there is anecdotal evidence that they may inform about language prediction as well. Furthermore, anteriorly distributed sentence-level negativities observed when processing syntactically difficult sentences have been previously tied to predictive mechanisms, whereby activated information would be temporarily stored in WM until it can be used later in time. Finally, in the domain of neural oscillations, alpha and/or beta power decreases have been linked with anticipation and prediction across different cognitive domains. As I will argue next, these signatures provide a starting point to seek signatures of anticipatory activity during predictive language processing.

Research aims

4 Research aims

The astounding efficiency with which humans comprehend language may be afforded by the generation of predictions that guide and speed up processing. As reviewed in the Introduction, there is now evidence that, at least under some circumstances, readers and listeners draw upon contextual information and stored knowledge to predict different aspects of upcoming linguistic input. Most evidence has been derived from facilitated processing of expected words, as indexed, for example, by reduced amplitudes of the N400 component (for a review, see Kutas & Federmeier, (2011). However, from the lens of predictive processing, relevant mechanisms should start to be mobilized earlier in time, and before external input confirms or disconfirms the prediction.

Based on this background, the goal of the present dissertation was to **investigate electrophysiological signatures of semantic anticipation in sentence comprehension.** To this end, we developed and tested a novel experimental paradigm to unveil neural activity in the prelude of critical words during sentence comprehension. From this starting point, the specific research aims of each study were the following:

The first aim was to **(1) investigate slow negative potentials (SNPs) as indices of semantic anticipation in spoken sentence comprehension (Study 1).** SNPs, such as the SPN, have qualified as reliable neural signatures of anticipation using a variety of tasks (Brunia et al., 2011), and there is anecdotal evidence of SNPs associated with context-driven expectancy in sentence comprehension (van Petten & Kutas, 1991; Besson et al., 1997). Joining these observations, we hypothesized that SNPs may serve as a reliable ERP signatures of prediction in sentence comprehension. To address this, we recorded ERPs while participants listened to sentences that afforded strong predictions, weak predictions or no predictions about their final word. Crucially, to elicit anticipatory responses, the final word appeared after a short interval (1 second). To obtain the full picture, we additionally assessed effects at the processing stage of words. To evaluate N400 context effects, the final word could be semantically expected or unexpected. This was also a necessary control to convincingly draw a link between anticipatory activity and linguistic prediction – a relative reduction of the N400 amplitude should be observed for expected words, indexing facilitated processing (Kutas & Federmeier, 2011).

The second aim of the thesis was to **(2) investigate SNPs as indices of semantic anticipation in reading sentence comprehension and track anticipatory activity over the course of sentence processing (Study 2)**. In Study 1, we confirmed the central hypothesis that SNPs would capture differences as a function of contextual constraint. To further understand the functional basis of this effect, we adapted the experimental paradigm for visual presentation, allowing to determine it was dependent on the modality of comprehension. In addition, we investigated slow activity dynamics by tracking word-by-word changes during sentence processing. This approach can unveil cumulative, sentence-level processes that are independent of the more punctate responses triggered by single events (Kutas & King, 1996), such as WM demands incurred by the active maintenance of information during sentence comprehension (e.g., Fiebach et al., 2002).

While the previous research aims were pursued by analyzing ERPs, the third aim of the thesis was to **(3) explore alpha and beta oscillatory activity as indices of prediction in sentence comprehension (Study 3).** Unlike ERPs, oscillatory activity is not necessarily phase-locked to an event, and thus can reveal additional and complementary information. In other cognitive domains, pre-stimulus alpha and beta decreases have been associated with mechanisms involved in sensory prediction (Arnal & Giraud, 2012; Jensen et al., 2012) and the encoding and retrieval of information (Hanslmayr et al., 2012). In contrast, little is known about the oscillatory mechanisms that support prediction in language. To contribute to filling this gap of knowledge, we performed a conjoint analysis of Study 1 and Study 2 to investigate the involvement of alpha and beta oscillatory activity during prediction in sentence comprehension, and to what extent the mechanisms supporting linguistic prediction are dependent on the modality of comprehension (reading and spoken comprehension).

Finally, after testing it with healthy adult population, as a clinical application of the experimental task, we used it to **(4) study predictive language comprehension in Parkinson's disease (PD) (Study 4)**. Besides motor impairment, PD also entails cognitive deficits that can progress in severity and meet the criteria for mild cognitive impairment (MCI) and dementia (Aarsland et al., 2010). Such deficits extend to language comprehension as well (for a review, see Pell & Monetta, 2008) and could be potentially related to difficulties in predictive language processing. In healthy population, older adults show diminished or absent ERP effects associated with prediction, suggesting that age-related cognitive decline may affect predictive language processing (see Section 3.1.4). Following this, we hypothesized that individuals with PD may have difficulties with linguistic prediction, which could lead to less efficient language processing in the long run. We first tested this **in individuals with PD with preserved cognition** and then also **in a sample of individuals with PD and MCI**, a diagnostic entity that has received a lot of attention in recent years to search for neural markers that can predict the development of dementia in PD. Finally, we also **tested the correlation between ERP signatures and verbal fluency scores,** since this neuropsychological measure has been previously associated with language prediction.

Study 1

Electrophysiological correlates of semantic anticipation during spoken comprehension

This study corresponds to:

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

5 Study 1

Electrophysiological correlates of semantic anticipation during spoken comprehension

5.1 Introduction

The complexity and vastness of human language contrasts with the seemingly easiness of interlocutors to understand and react to linguistic utterances. The solution to this paradox might lie in the brain's ability to predict upcoming events and prepare for their occurrence (Bar, 2007). From predictive-based models of language processing, the concept of prediction refers to the pre-activation of specific concepts or their features before they are perceived (Kutas, DeLong, & Smith, 2011). From this perspective, incoming contextual information and prior knowledge are interactively combined to guide the pre-activation of the most probable continuations to the unfolding speech, which might explain why words that are more predictable are read faster, more likely to be skipped when reading (Ehlrich & Rayner, 1981; McDonald & Shillcock, 2003) and better decoded under circumstances of degraded speech (Clos et al., 2012; Miller et al., 1951).

Electrophysiological evidence coming from ERP studies of the N400 component has been a very important tool for prediction-based theories in language. The N400 (Kutas & Hillyard, 1980) is a negative-going voltage deflection peaking approximately 400 ms after the onset of any potentially meaningful word (Kutas & Hillyard, 1984). On experimental grounds, the predictability of words is usually operationalized as cloze-probability– the percentage of individuals that supply that word as a continuation to a particular sentence (Taylor, 1953). The cloze-probability of a word depends on the degree of constraint of its preceding context. Highly constraining contexts typically have a best completion with a much higher cloze-probability than any other continuation (e.g. "The dentist proceeded to clean her... teeth") while low constraining contexts have more than one likely continuation (e.g., "The meeting was arranged for the... morning/afternoon/evening") and their cloze-probabilities are lower. Importantly, the amplitude of the N400 follows a graded function that is negatively correlated with the cloze-probability –i.e. predictability– of the eliciting word given the preceding context (DeLong, Urbach, & Kutas, 2005; Federmeier, Wlotko, Ochoa-Dewald, & Kutas, 2007; Kutas & Federmeier, 2011; Kutas & Hillyard, 1984; Van Petten, Coulson, Rubin, Plante, & Parks, 1999;). From the perspective of active prediction models, the N400 amplitude reduction to words with higher cloze-probabilities is an evidence of word processing facilitation as a consequence of successful word pre-activation. However, it is also possible to explain the reduced N400 amplitude in the absence of predictive processes. Other authors have argued that it is more effective to passively wait until words are correctly identified rather than to anticipate them, given that in language there are usually an infinite number of potential continuations to the same sentence, making prediction an unviable strategy (Jackendoff, 2002; Morris, 2006). Language would be processed in a bottom-up, stimulus-driven fashion (see Altmann & Steedman, 1988; Marslen-Wilson & Tyler, 1980; Van Berkum, Brown, & Hagoort, 1999 for discussions) and the context and other top-down influences would exert their influence only after complete word identification. From the perspective of passive integration models, the N400 modulation would merely reflect word integration processing costs upon their receipt, reduced for eliciting words that fit the prior context-based information better. This is referred to as the active prediction versus passive integration dilemma.

A way to disentangle the contribution of prediction and integration processes is to provide evidence of prediction preceding the identification point of the word that is being predicted (Kutas et al., 2011). In line with this idea, DeLong et al. (2005) used the fact that in English, the a/an indefinite articles are functionally and semantically identical and are alternated only as a function of whether the initial phoneme of the noun they precede is a consonant (calling for "a") or a vowel (calling for "an"). Contrary to passive integration theories, they found that indefinite articles that mismatched the expected following word elicited larger N400 responses, suggesting that individuals were making predictions about words and that those predictions were fulfilled or violated as soon as the article was encountered, that is, before the actual word was perceived (see also, for variants of this paradigm, Van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005; Wicha, Bates, Moreno, & Kutas, 2003). Similarly, studies of the time-course of spoken word identification show that the N400 congruity effect (divergence in the ERP between congruent and incongruent words) begins 200 ms prior the eliciting word’s isolation point – before it differs completely from any other congruent candidate word that starts in the same way (Van Petten et al., 1999). Also, in a recent magnetoencephalograpy experiment (Dikker and Pylkkänen, 2013), picture primes were used to manipulate contextual constraint. The pictures did or did not allow predicting specific nouns (i.e., the picture of an apple predicted the very same concept, but a picture of a grocery bag was predictive of any fruit). After the prime, the word “the” followed by a noun was presented, that could match or mismatch the prediction. They found that the primes that were predictive triggered enhanced activation of a top-down network that they interpreted as an evidence of word preactivation at different levels.

All these results converge on the idea that language processing is not strictly stimulus-driven. Instead, individuals actively use the contextual information and prior knowledge to prepare for the upcoming words. From prediction-based models, the facilitation of candidates that hold semantic and functional similarities with the most expected word suggests that the pre-activation process is strongly influenced by how information is stored in long-term memory (Kutas & Federmeier, 2000). However, the type of content that is being pre-activated is still a matter of investigation, whether the pre-activation is the complete representation of the predicted item that is pre-activated or some of its associated features (e.g., semantic or morphosyntactic features) (Brothers, Swaab, & Traxler, 2015; Huettig, 2015, for a review; Lau, Holcomb, & Kuperberg, 2013).

Certainly, the aforementioned N400 component studies argue strongly in favor of prediction in language. However, the N400 component is observed when the target word has already been presented, and therefore it is a correlate of the word processing. Because of this, it does not provide direct information about the brain correlates of the anticipatory processes that might be taking place before the target word is presented. To the best of our knowledge, to date there is only one study that might have provided some evidence of semantically related anticipatory processes previous to the target word presentation in the language domain. Besson, Faita, Czternasty and Kutas (1997) reported two experiments in which high semantically constraining sentences (proverbs) and low semantically constraining sentences were used. In half of the trials, an unexpected 600 ms pause was inserted between the penultimate and the ultimate word of the sentence. In their first experiment they used visual stimuli presented at a slow pace (200 ms each, with a 500 ms SOA). They found that, during the pause, and therefore previously to the presentation of the final word, a Contingent Negative Variation (CNV) developed. This CNV had a higher amplitude in the low constraining sentences than in the high constraining sentences, this is, the former had more negative voltages than the latter. According to the authors, the rather slow rate of word presentation in the visual modality could have provided participants with sufficient time to anticipate the final word much before its occurrence so the CNV reflected the amount of expectancy towards the final word, more positive for more predictable continuations.

In their second experiment, Besson and colleagues (1997) used natural speech auditory stimuli. In this case, when they analyzed the results of the pause period, they also observed more negative voltages in the low constraining sentences than in the high constraining sentences. However, in this case instead of the sustained CNV observed in their first experiment, they obtained a marked emitted potential. According to the authors, the pause was more salient in the auditory modality because it strongly disrupted the temporal cadence of natural speech. Therefore, the emitted potential reflected surprise to the sudden interruption. This surprise would have been larger for high constraining sentences, where highly expected continuations were suddenly unfulfilled than in the case of the low constraining sentences.

Slow negative potentials, such as the CNV, have been systematically observed in other domains and are known to reflect anticipatory attention for upcoming relevant events (Brunia & Van Voxtel, 2001; Brunia, Van Boxtel & Böcker, 2011; Brunia, Hackley, van Boxtel, Kotani & Ohgami, 2011). In particular, the CNV (Walter, Cooper, Aldridge, McCallum, & Winter, 1964) is a negative slow brain potential, also known as the "expectancy wave", that shows up if a warning signal announces that, imminently, another stimulus will arrive, requiring some response. The CNV can be subdivided into two phases (Connor & Lang, 1969), an early phase (CNVe) that immediately follows the warning signal, and the terminal phase (CNVt) that is comprised by the Readiness Potential (RP) and the Stimulus Preceding Negativity (SPN) (Brunia et al., 2011; Van Boxtel & Brunia, 1994). The SPN is another slow ERP component that progressively increases in amplitude as subjects are waiting for a stimulus that provides relevant information, such as performance feedback, instructions or affective stimuli (Van Boxtel & Böcker, 2004). Also, in learning paradigms, the amplitude of the SPN varies as learning advances (Morís, Luque & Rodríguez-Fornells, 2013) –that is, as future events become more predictable. In this experiment the voltage of the SPN became more positive as the incoming feedback became more predictable. Given all this, and the results obtained by Besson and colleagues' (1997) results, slow brain potentials are good candidates for direct correlates of anticipatory processes in language comprehension.

The goal of the present study is to investigate if a slow component-like can be consistently observed before word perception and, if this is the case, to determine to what extent it might be a correlate of context-based word anticipation. To do so, we used an auditory delay paradigm. We presented participants semantic contexts of varying semantic constraint. Semantic contexts were either high (HC) or low (LC) constraining, and we included a non-semantic condition (NS) that did not provide a meaningful context. Also, a 1000 ms silent delay was inserted prior to the final word of each sentence. As have already described, a similar study reported the development of a slow potential in the visual but not in the auditory modality (Besson et al., 1997). Nevertheless, we hypothesized that, with an appropriate control of the variables detailed hereafter, a component reflecting the proposed anticipatory process should be observed regardless of the sensory modality. Crucially, we presented the delay systematically in all trials, in contrast with the procedure used by Besson et al. (1997), in which only half of the sentences had a delay. By presenting this delay in all trials, we removed surprise effects due to an unexpected speech disruption. Another difference was that, in contrast with Besson et al. (1997), all of the sentences employed were novel constructions, controlling the differences in the familiarity of the sentences.

As a consequence of controlling the aforementioned variables, any difference in the electrophysiological response observed during this delay should be exclusively due to differences in the contextual constraint of the sentence, without any of the confounds that were present in previous experiments. We hypothesized that a slow potential reflecting anticipation would develop during the delay, and that its amplitude should vary proportionally to how much information the context provided about the upcoming word. That is, its amplitude should be modulated as a function of the contextual constraint of the sentence. If we observe a modulation of an ERP component during the delay that develops prior to the target word and that is sensitive to the semantic constraint of the sentence, it would be an excellent candidate for a correlate of semantic anticipation, while providing additional support for prediction-based models in language processing. Additionally, we manipulated the congruency of final words in the HC and LC conditions, and measured the N400 component that they elicited.

5.2 Methods

Participants

Twenty-two right-handed young adults (12 females, M_{age} = 21 years, *SD* = 1.8, age range = 18–24), were paid to participate in the experiment after giving written consent. None of the participants reported health problems or previous neurological disorders. Participants had normal hearing and normal or corrected-to-normal vision. All participants were native Spanish speakers and were on the course of completing or had already completed a higher-level education degree. The study was approved by the Ethics committee of the Hospital Universitari de Bellvitge.

Procedure

Participants were comfortably seated in front of a computer screen placed 70 cm away from them. After electrode application, headphones were placed over the cap. Participants were informed that they would listen to sentences disrupted by a short pause before the final word. They were told to pay special attention to the final word because they would have to complete a memory recognition test at the end of each block. Finally, they were briefed on the importance of minimizing movement and were told to synchronize their blinks to the blinking signal presented on the screen during the experiment completion and to prevent them at other times.

Stimuli and Task

We designed an auditory delay paradigm in which participants listened to sentences of a varying degree of contextual constraint that had a 1 s silent pause before the presentation of the last word. A total of 392 sentences (176 highly constraining sentences, 176 low constraining sentences and 40 non-semantic sentences) and 392 final words were used as stimuli. Of those words, 176 were congruent final words –best completion of the sentence–, 176 were incongruent final words, and 40 were neutral words for the non-semantic condition. We used the materials from Mestres-Missé, Rodríguez-Fornells and Münte (2007), who created and categorized sentences in Spanish, that were either low constraining (mean cloze probability 6.1 ± 10.3% standard deviation), or highly constraining (mean cloze probability 76.0 ± 17.7% standard deviation). The final words to the sentences were always nouns. For the non-semantic condition (NS) condition, we randomly picked 40 sentences from the low contextual constraint condition and scattered the vowels of each word so that they were completely unintelligible but still preserved the grammatical structure. The incongruent final words were obtained from the ESPAL database (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013), a Spanish corpus of lexical stimuli, matching the congruent words in length, number of syllables and frequency. All the linguistic stimuli were transformed into audio using a voice-synthesizer software (Loquendo TTS Director, 2005). This software creates natural-sounding audio, very close to natural speech, while allowing precise control of the speech rate, the amplitude and the prosody of the speech (see Supplementary materials for a sample sentence of each condition). After audio conversion, the first part of the sentences had a mean duration of 1966 ms (*SD* = 288) with similar means in each condition (HC = 2000 ms, LC = 1925 ms, NS = 1992 ms). The final words had a mean duration of 490 ms (*SD* = 122) with similar durations across conditions (mean durations: Congruent = 485 ms, Incongruent = 494 ms, Neutral = 495 ms). The experiment was divided into 10 blocks of 20 sentences in which 20% were non-semantic sentences, 40% highly constraining sentences and 40% low constraining sentences (within the conditions of sentences with semantic content, half ended with a congruent word and half with and incongruent word). Each block lasted approximately 2.5 minutes (*M* = 8456 ms). Within each block, the order of the sentences was random.

Table 1. An example of the four potential combinations of sentence type and final word type (1 to 4) and an example of a non-semantic sentence type (5). High and low-cloze probability conditions were counterbalanced because each final word (either congruent or incongruent) fell both into the high or low-cloze category depending on which sentence type they randomly followed.

	Example Sentences	*Constraint-Congruence Combination*
1	I have never flown in a *plane*	High constraint-Congruent
2	The dot in the sky must be a *plane*	Low constraint-Congruent
3	I have never flown on a *drink*	High constraint-Incongruent
4	The dot on the sky must be a *drink*	Low Constraint-Incongruent
5	Tehtod no tehykstusmeb a *trip*	Non-Semantic

The task proceeded as follows (**Fig. 1**). A fixation point (a cross) remained on the screen from the beginning until the end of the trial. Two seconds after the presentation of the fixation point, participants listened to a sentence that included a 1 second delay before the presentation of the final word. Then, they waited for 1 second until the blinking signal appeared (a picture of an eye presented at the center of the screen). The blinking signal remained for 2 seconds before the next trial began. With the objective of ensuring that participants paid attention to the relevant stimuli, a memory recognition test followed each block. In this test, ten words were presented visually one at a time; half were final words previously presented and half completely new ones. Participants had to respond, pressing a key, whether they had heard the word in the block or not. Before the start of the task, an instruction screen indicated the correspondence between "z" and "m" keys and "yes" and "no" answers, counterbalanced across participants. Once the memory recognition task began, each word remained on the screen until an answer was supplied, followed by a 600 ms fixation cross until the next word appeared. At the end of the memory task the next block started right away, except for even numbered blocks, which were followed by a pause that could be resumed anytime.

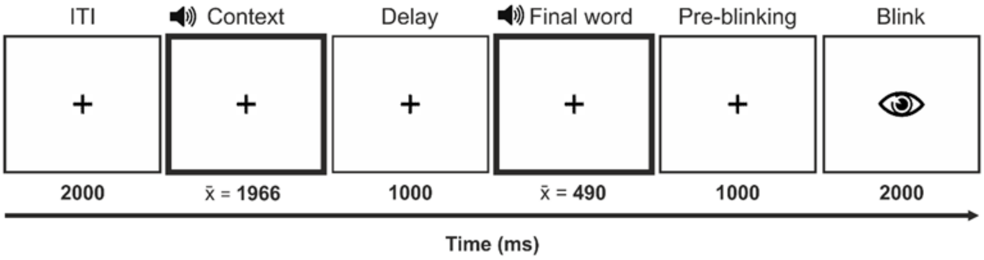

Figure 1. Step-by-step depiction of a trial. In each trial, participants listened to a sentence that had a 1000 ms silent delay before the presentation of the final word. After the presentation of the final word, a signal announced participants that they could blink. Each block included 20 trials and each participant completed a total of 10 blocks.

For the purpose of this study, we were particularly interested in controlling for potential confounds in the contextual constraint manipulation. To do that, each HC sentence was paired with a LC sentence that shared the same congruent and incongruent final words (for instance, both the highly constraining sentence "I have never flown on a…" and the low constraining sentence "The dot in the sky must be a…" shared "plane" as their congruent ending and "drink" as their incongruent). Therefore, across subjects, final words did not change their congruence category ("plane" was always congruent, and "drink" was always incongruent) but they changed their cloze-probability value depending on which sentence type (HC or LC) they randomly followed. Within subjects, the combination of a particular sentence type and final word type was randomly picked in each trial, and once a specific combination was presented (i.e., "I have never flown on a… *plane*"), the other three remaining potential combinations of the same pair were excluded from selection for the same participant. At the end, the high and low constraint effects were completely counterbalanced across subjects because every final word was equally used as a low and high-cloze ending. Moreover, although the congruence categories remained invariant across subjects, every pair of congruent and incongruent words (i.e. "plane" and "drink") were matched in word length, word frequency, familiarity, imaginability and concreteness.

In the case of the NS condition, each final word was matched with the corresponding congruent and incongruent pair of its scattered sentence. For example, if the sentence used for scattering was "The dot on the sky must be a…" and final word pair was "plane" and "drink", then final word "trip", matched in the variables mentioned before, was used for the non-semantic condition.

Electrophysiological recording

The electroencephalographic signal (EEG) was recorded from 31 scalp electrodes placed at standard positions of the 10–20 system (electrode positions: Fp1/2, Fz,F3/4, F7/8, FCz, FC1/2, FC5/6, Cz, C3/4, T3/4, T5/6,CP1/2, CP5/6, Pz, P3/4, PO1/2, Oz, left and right mastoids) mounted in an elastic cap. Two electrodes placed below the right eye and 1 cm from its outer canthus registered respectively the horizontal and vertical electro-oculograms (EOG). Electrode impedances were kept under 3 KΩ when possible and always below 5 KΩ. The EEG was amplified with BrainAmps amplifiers (BrainProducts, München) with an online band-pass filtering of 0.015-1000 Hz and sampled with a frequency of 1000 Hz.

ERP Data Analysis

Data were re-referenced off-line to the average of the mastoid electrodes. Before analysis, the data were filtered using 50 Hz notch filter (to attenuate electrical line noise) and a 30 Hz low-pass Butterworth filter, with a roll-off of 12 dB/oct, as implemented in the ERPLAB toolbox V4.0 (López-Calderón & Luck, 2014). An additional 8Hz filter was applied to the figures presented. We applied Mestres-Missé and colleagues (2007) criteria for trial rejection: all epochs with an activity over ± 85 µV in the ocular channel or ± 200 µV in any other channel were removed. Furthermore, after visual inspection, trials contaminated by blinks, excessive muscle activity or big drifts were eliminated. The percentage of trials included after the rejection was not significantly different between conditions of contextual constraint (HC, M = 93.9 %; LC, M = 95.2 %; NS, M = 94.77 %; $F(2,42) = 1.031$, $p = .37$, $\eta^2_p = .04$) or between conditions of the final words (HC congruent, M = 96.2 %, HC incongruent, M = 96.4 %; LC congruent, M = 97.4 %; LC congruent, 96.4 %; Neutral = 96.6 %; $F(4,84) = .881$, $p = .48$, $\eta^2_p = .04$).

The analysis was carried out in two different parts. The first was focused on the activity in the delay period between the initial portion of the sentence and its final word. The epochs for this analysis were time-locked to the onset of the delay, ranging from -100 to 1000 ms. The 100 ms pre-delay interval served as baseline. For this analysis, the target time window was the final 200 ms of the delay, of the fronto-central electrodes. This selection was made a priori and was based directly on the results and analysis protocol of Morís et al. (2013). They showed in a similar paradigm, although non-linguistic, that changes in SPN increase over time, and that the effect is maximal right before the presentation of the expected stimulus, as it had already been described in the literature (see Brunia et al., 2011 for a review). To provide additional evidence of the robustness of the effect, another set of analyses was carried out using a baseline going from 0 to 100 ms post-onset of the delay (see Supplementary Materials).

The second analysis, corresponding to the N400, was time-locked to the onset of the final word and included the time window from -100 to 1000 ms, the baseline being the 100 ms pre-stimulus. In this case, we used standard N400 parameters. The time window selected ranged from 300 to 500 ms, and we analyzed the Pz electrode (e.g., Mestres-Missé et al., 2007). An additional analysis of the N400 component using a cluster of central-posterior electrodes was also carried out (see Supplementary Materials). After baseline correction, we computed the average for each subject and condition for each analysis separately. Statistical analyses were run with SPSS 21. All the Linear Mixed Model analyses were run using an autorregressive matrix of covariance.

5.3 Results

ERPs during the delay

A first negative potential peaked approximately at 160 ms after the onset of the delay, followed by a slow potential, starting approximately at 200 ms, which developed from that moment to the end of the delay (see **Fig. 2**).

A Linear Mixed Model analysis was conducted on the mean amplitudes of six frontal and fronto-central electrodes, on right (F8 and FC6) central (Fz and FCz) and left (F7 and FC5) locations within the 800-1000 ms time window after the onset of the delay. We included three factors: Condition (3 levels, HC, LC and NS), Laterality (3 levels, right, central and left) and Position (2 levels, frontal and fronto-central). The model yielded a significant effect of Condition ($F(2, 248.62) = 41.657, p < .001$). Post hoc t tests for related samples showed significant differences between the three conditions ($p < 0.05$ in all the comparisons). The size of voltage amplitudes increased in a scaled fashion with HC as the most negative condition (M = 0.52 μV), followed by LC (M = 1.14

μV) and NS as the most positive (*M* = 2.09 μV). Specifically, the mean voltage of HC was significantly smaller than the mean voltages of the LC condition (*p* = .001, 95% CI of the difference [0.28, 0.95]), and of the NS condition (*p* < .001, 95% CI of the difference [1.23, 1.91]). At the same time, the mean of LC was smaller than the mean of NS (*p* < 0.001, 95% CI of the difference [0.623, 1.29]). The Laterality main factor was also significant (*F*(2, 119.17) = 11.71, *p* < .001), but not the Position factor, their interaction, or any of the interactions involving the Condition factor (p > .182 in all of the cases)

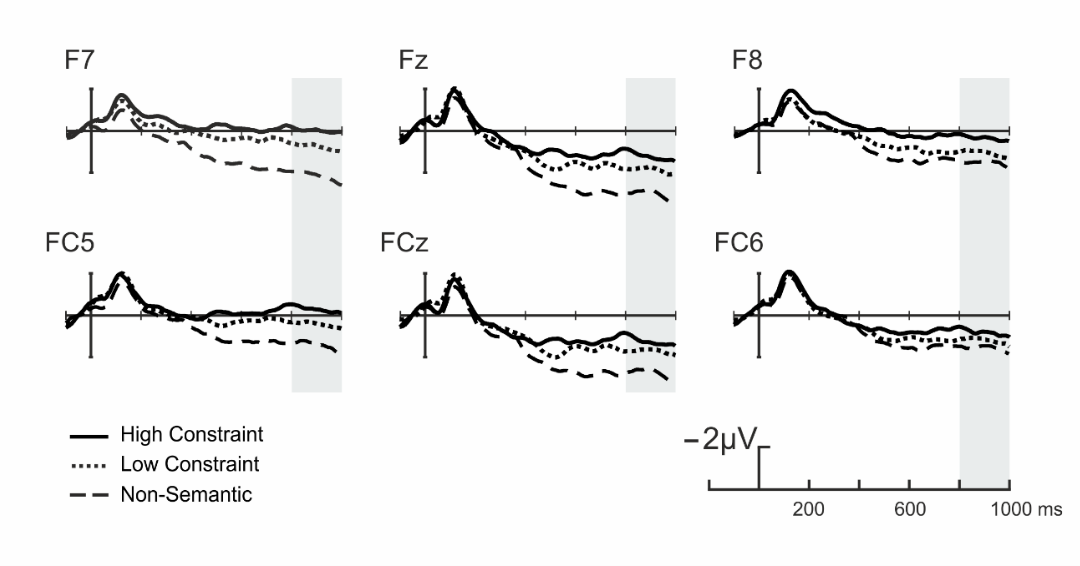

Figure 2. Grand averages at electrodes with left (F7, F5), central (Fz, FCz) and right (F8, FC6) locations showing the slow potential increasing along the interval and the changes in amplitude among conditions (HC, LC and NS). Negative is plotted upward. A slow increasing potential developed along the delay interval before the presentation of the final word. Following Morís et al. (2013) statistical analyses were conducted on the 800–1000 ms time window, shadowed in grey in the figure.

The scalp distribution of the differences between conditions, for the same 800–1000 ms time window is depicted on **Fig. 3.** The topographic maps of the bottom row show the differences between mean voltages of HC and NS, with a frontal and fronto-central distribution and a tendency to show larger amplitude in the left hemisphere. The topographic maps of the top row represent the mean voltage differences of HC and LC, smaller but still significant. They are also maximal at left frontal and fronto-central sites. As mentioned before, voltage becomes more positive as the sentence's cloze-probability decreases, being the most negative for HC and most positive for NS.

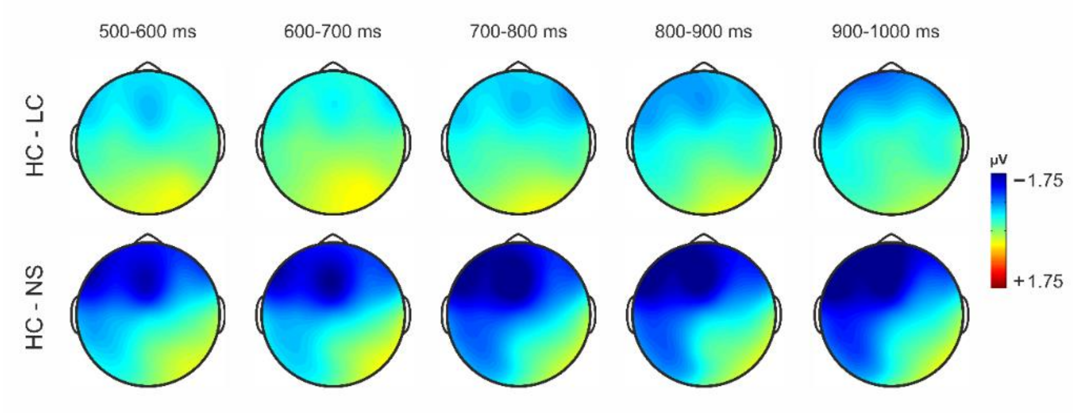

Figure 3. Topographical maps of the mean voltage differences between HC and LC (top) and between HC and NS (bottom) for the mentioned time window. The voltage values for the maps range ± 1.75 µV.

N400 effects

A Linear Mixed Model was performed on the Pz electrode using the time window of 300 to 500 ms from the onset of the final word (See **Fig. 4**) (Mestres-Missé et al., 2007), with Contextual Constraint (2, High and Low) and Congruency (2, Congruent and Incongruent) as factors. The analysis revealed significant main effects of Contextual Constraint ($F(1,83.334) = 15.21$, $p < .001$) and Congruency ($F(1,56.77)= 48.54$, $p < .001$). The Contextual Constraint × Congruency interaction was also significant ($F(1,60.92) = 5.51$, $p = .022$) (see Fig. 4, left). Post-hoc tests revealed that incongruent endings to the HC condition produced larger amplitudes ($M = –3.6$) than congruent endings ($M = –0.62$) ($p <.001$, 95% CI of the difference [2.05 3.92]), just as the amplitude of incongruent endings to the LC condition ($M = –4.49$) was larger than that of congruent endings ($M = –3.16$) ($p = .006$, 95% CI of the difference [0.40 2.27]). Within congruent endings, the mean voltage of HC congruent was significantly more positive than LC congruent ($p < .001$, 95% CI of the difference [1.42 3.66]), while LC incongruent was not significantly different from the HC incongruent ($p = .117$, 95% CI of the difference [-0.23 2.01]). An additional analysis using a cluster of four electrodes (Cz, CP1, CP2 and Pz) found the same pattern of results (see Supplementary Material)

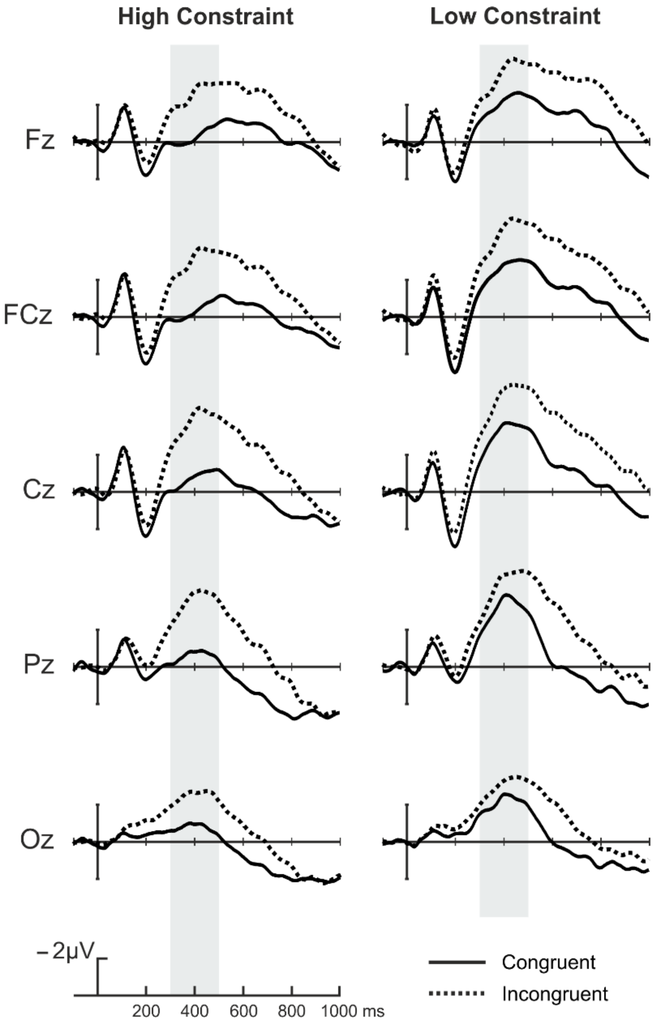

Figure 4. Grand averaged ERP waveforms in the 1000 ms following the offset of the final word (All midline electrodes, Fz, FCz, Cz, Pz, Oz) for all semantic conditions. The grey area corresponds to the 300-500 ms time window where the N400 effect was analysed. A negativity (N400) was elicited, especially on parietal and central sites.

The incongruent minus congruent difference waveforms were obtained for both Contextual Constraint conditions (see **Fig. 5**, left panel) on the Pz electrode. Onset latency was determined and subjected to statistical analyses. Following Rodríguez-Fornells, Kurzbuch, &, Münte (2002) procedure, onset latencies were calculated via a stepwise series of *t*-tests (step size = 4 ms). For each test, data from a time-window of 50 ms was averaged (i.e., point of measure ± 25 ms) in the full 1000-ms period after the final word onset. Onset latency was defined as the point at which four consecutive *t*-tests showed a significant difference from zero ($p < 0.05$). The difference wave for HC endings diverged from baseline from 265 to 877 ms (for a 612 ms duration, $2.69 < t(21) < 5.54$) and the difference wave for LC endings diverged from baseline from 346 to 817 ms (a duration of 471 ms, $2.06 < t(21) < 7.04$). A comparison between the difference waveforms of the two conditions revealed significant differences between 316 to 348 ms and then from 369 to 458 ms (a total duration of 89 ms, $2.07 < t(21) < 2.49$) (see **Fig. 5**., right panel).

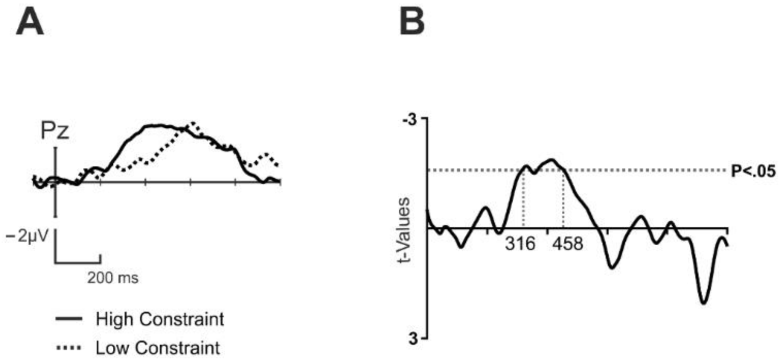

Figure 5. Left panel. Plot of the significant interaction Cloze-Probability x Congruency. Right panel. Grand averaged ERPs for the Pz electrode elicited on congruent endings in HC and LC conditions, along with the NS condition. A smaller N400 effect is associated to HC congruent endings compared to LC congruent and NS endings.

The topographical mapping of the temporal evolution of the congruency effect also evidences the latency difference (see **Fig. 6**) and reveals a typical centro-parietal distribution of the N400 (Kutas & Hillyard, 1984).

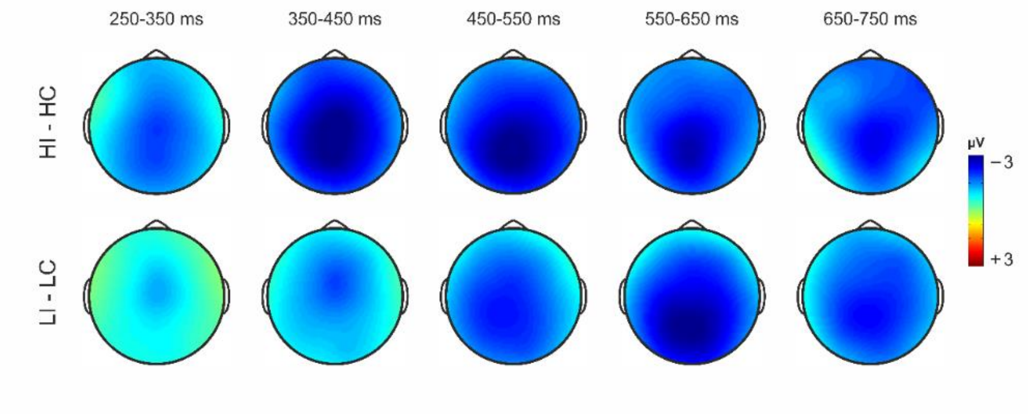

Figure 6. Temporal evolution of the congruency effect for high (top row) and low (bottom row) cloze-probability endings. Spatial distribution of the differential voltage (incongruent - congruent) every 100 ms in the 250- to 750-ms time window. The topographical mapping evidences the earlier peak of the effect for high compared to low cloze-probability endings. The range of voltage values for the maps is ± 3 μV.

5.4 Discussion

This study investigated correlates of prediction-related processes during speech processing taking place before the presentation of the target word. To accomplish this goal, we inserted a constant 1 s delay between the penultimate and the final word of a sentence. The semantic context of the sentences varied in the degree of contextual constraint, establishing high, low or no expectancy towards the final word, that could turn out to be congruent –the best completion– or incongruent. The ERP analysis during the delay period unveiled a slow potential, with an amplitude sensitive to the level of contextual constraint, being more negative as contextual constraint increased. After the presentation of the final word of the sentence, we observed a canonical N400 modulation to semantic fit and contextual constraint and we report a delay in the onset of the N400 effect for low levels of contextual constraint.

First, we focus on the main similarities and differences between the slow potential reported and the SPN. The SPN is typically elicited in motivation and learning paradigms whenever a delay is introduced prior to the occurrence of motivationally relevant stimuli (e.g., monetary rewards, performance feedback, evocative photos, or painful stimuli) which are expected (Brunia et al., 2011; Damen & Brunia, 1987; Kotani et al., 2015). In these studies, the SPN is interpreted as an index of expectancy, as its amplitude usually becomes larger with increasing levels of expectancy or predictableness of the upcoming stimuli (Kotani et al., 2003; Morís et al. 2013). Similar to the SPN, the amplitude of the slow potential reported here follows a graded order consistent with the differences in the level of expectancy that each condition establishes toward the final words, as well as a progressive amplitude increase in time (Walter et al., 1964; Brunia et al., 1988) and a mainly fronto-central distribution (e.g., Mattox et al., 2006; Morís et al., 2013; Hackley et al., 2014). However, the most notable difference is that the voltage pattern is the inverse compared to the typical SPN, as in the slow potential we report the amplitude is more negative for expected stimuli, instead of more positive. In that respect, most descriptions of the SPN to date come from tasks that use non-linguistic stimuli and that might entail distinct cognitive processes (see Brunia et al., 2011, for a review) so the activation of a different cortical generator might explain this difference (Luck, 2012). The use of linguistic material with semantic content might also explain the apparent left-hemisphere dominance on fronto-central sites of the slow potential we observed, which differs from the SPN right-hemisphere preponderance sometimes observed (Damen & Brunia, 1987; Ohgami et al., 2004).

While the use of linguistic materials might be a significant factor in the direction of the effect, the order of the voltage pattern cannot be attributed solely to this because it also differs from some results in language-related studies. In fact, in the domain of language, the available evidence on slow potentials is rather inconsistent. On the one hand, in line with our results, Kaan and Carlisle (2014) found a slow potential component with more negative amplitudes during the delay to predictive compared to random letter sequences, in a visual paradigm. On the other hand, as previously mentioned, a similar study by Besson and colleagues (1997) did not observe the development of a slow component in the auditory modality. We think that the experimental difference that might have allowed for the observation of the slow component in our case was the elimination of the surprise effect to the delay by introducing it in all the trials, instead of introducing it randomly and only in half of them. Then, in the visual modality, Besson and colleagues' found a slow potential but they encountered more negative amplitudes for sentences with a lower contextual constraint –that is, opposite to our findings–. One possibility is that the differences in the experimental paradigm –the type of manipulation, the nature of the anticipated stimuli, the duration of the anticipation interval, among others– activated different cortical generators of the slow potential (Kotani et al., 2015; Stravopoulos & Carver, 2014; Van der Molen et al., 2013) that might account for this. A first critical difference between the two studies is that our task was auditory instead of visual, although we already mentioned one study in the visual modality that converge with our observations (Kaan & Carlisle, 2014). Secondly, Besson and colleagues' used proverbs for the condition of high contextual constraint while we used novel sentences. Proverbs are fixed expressions in memory which might engage qualitatively different pre-activation processes compared to previously unheard sentences (Cacciari & Tabossi, 1988; Vespignani, Canal, Molinaro, Fonda & Cacciari, 2010) which represent a more common situation and are the type of sentences that we used. On that point, we think that the materials that we used in our experiment might be better suited in order to reveal how contextual information impacts natural language processing.

Altogether, comparable evidences are scarce and rather inconsistent in the domain of language. Despite this, in a more general sense, our finding matches with previous descriptions of slow event-related potential modulations appearing in between relevant linguistic stimuli (Besson et al., 1997; Fiebach, Schlesewsky, & Friederici, 2002; Kaan & Carlisle, 2014; Kutas & Hillyard, 1980, 1984; Kutas et al., 1988) or across the course of a whole sentence (Münte, Schiltz, & Kutas, 1998; Nieuwland, 2015; Van Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991). The generally convergent observations in these ERP studies provide soundness to the interpretation that the slow potential we describe here might reflect some language-related process. In particular, given their similarities, we consider that it cannot be ruled out that the component we report here might be a modulation of the SPN. Further investigation is required to investigate to which extent the slow potential activity recorded here might reflect an expectancy-based process ubiquitous to other domains or a language-specific process.

Our results are in line with models of anticipatory language processing that state that contextual information is *not only* processed incrementally as the information unfolds in time but that the brain proactively uses top-down information to pre-activate upcoming words, or their associated features, and facilitate their processing upon receipt (Federmeier & Kutas, 1999; Kamide, 2008; Kutas et al., 2011 for a review, McClelland & Elman, 1986; Van Berkum et al. 2005; Van Petten et al. 1999; Van Petten & Luka, 2012). The amplitude modulation of the slow potential revealed that contextual constraint had an impact on cognitive processing even before the final word was perceived. In semantically meaningful conditions predictions would have been generated, producing different levels of expectancy –as a function of contextual constraint– towards the upcoming word. Consistently, we replicated the classical modulation of the amplitude of the N400 component to contextual constraint (Besson et al., 1997; Federmeier & Kutas, 1999; Federmeier, 2007; Kutas & Hillyard 1984) and to the congruency of the completions of the sentence (DeLong et al., 2005; Kutas & Hillyard, 1983; Kutas & Federmeier, 2000; Van Petten et al., 1999).

We found an earlier onset of the N400 effect for high compared to low contextual constraint. This result is consistent with the idea that as the contextual constraint increases, a larger number of lexical or/and semantic features associated to this specific context could be pre-activated (Boudewyn et al., 2015) and be straightforwardly compared with the actual bottom-up input, allowing to detect a mismatch sooner than in the case of low contextual constraint, where more than one lexical candidate might be pre-activated. In line with this result, Cermolacce, Scannella, Faugère, Vion-Dury, & Besson (2014) recently reported an N250 congruity effect for well-known familiar proverbs over unfamiliar sentences. In the case of proverbs, a perceptually-based representation of the specific upcoming word might be pre-activated causing an almost immediate perceptually-based mismatch when the input is incongruent. In our experiment we cannot attribute the onset difference to a congruency effect due to confound with cloze-probability effects (congruent and incongruent words have more similar cloze probabilities in low than in the high contextual constraint conditions). Despite this limitation, we think that the results in both studies are convergent and that they reinforce the interpretation of the N400 effect from prediction-based accounts.

A question that must be taken into account is whether the artificial introduction of a delay in sentences might lead to the engagement of processes that would not work in natural speech situations. The delay might provide an unusually long time to generate predictions that otherwise might not have been formulated (Rayner, Ashby, Pollatsek, and Reichle, 2004). It would be desirable for future experiments to explore whether anticipatory processes –such as the slow potential that we report– are also elicited under more naturalistic circumstances. Another interesting approach for future studies would be to explore the possible relationship between the processes indexed by the reported slow potential and the well-known N400 component.

Theoretically, if we interpret the reported slow potential as a correlate of prediction generation and the N400 component as an index of the level of mismatch between the predicted and the actual input, we would expect to find some degree of correlation between them. Our design did not allow us to test this possibility. Firstly, it would be necessary to use more than two levels of contextual constraint to allow matching a wider range of values of the two components. Secondly, in our task the critical time window of the slow potential overlaps partially with the baseline of the N400 component. This is not a problem as long as the two components are interpreted separately but it could produce an autoregression effect that would distort any direct correlation between them.

Regarding the main goal of the study, we found an electrophysiological correlate that shows a modulation consistent with the idea that forthcoming words or their associated semantic features might be pre-activated prior to their perception. Due to its shape and timing, the slow potential cannot be a correlate of the processing of the target word itself and it must reflect the operations of processes engaged before the presentation of such word. Given these features, the classical confound in prediction-based conclusions of N400 component studies –that is, that any difference in brain activity merely reflects context-driven facilitation in word integration– does not apply for the current data. While the slow potential might capture integration processes of the preceding context, it cannot reflect integration processes of the final, potentially predicted word. Overall, the present study contributes with novel evidence that is convergent with prediction-based interpretations drawn from previous N400 component studies. Altogether, both ways of approaching the investigation of processes involved in spoken language comprehension provide results in favour of anticipatory processing in language. Exploring the nature of the neural correlates of anticipation processes will continue to enrich our view of classical effects and theoretical frameworks that are currently used to explain them.

5.5 Supplementary materials

To provide additional evidence of the robustness of the effect, another set of analyses was carried out using a baseline going from 0 to 100 ms post-onset. The rest of the details are as described in the main text.

ERPs during the delay

A slow potential as that described in the main text of the manuscript was observed. A Linear Mixed Model analysis was conducted on the mean amplitudes of six frontal and fronto-central electrodes, on right (F7 and FC5), central (Fz and FCz) and left (F8 and FC6) locations within the 800-1000 ms time window after the onset of the delay. We included three factors: Condition (3 levels, HC, LC and NS), Laterality (3 levels, right, central and left) and Position (2 levels, frontal and fronto-central). The model yielded a significant effect of Condition (F(2, 278.04) = 40.94, p < .001). Post hoc t tests for related samples showed significant differences between the three conditions (p < 0.05 in all the comparisons). The size of voltage amplitudes increased in a scaled fashion with HC as the most negative condition (M = 1.1 μV), followed by LC (M = 1.88 μV) and NS as the most positive (M = 2.4 μV). The mean voltage of HC was significantly smaller than the mean voltages of the LC condition (p < .001, 95% CI of the difference [0.51, 1.05]), and of the NS condition (p < .001, 95% CI of the difference [1.02, 1.59]). At the same time, the mean of LC was smaller than the mean of NS (p < 0.001, 95% CI of the difference [0.25, 0.79]). The scalp distribution of the differences between conditions, for the same 800–1000 ms time window. The Laterality main factor was also significant (F(2, 107.12) = 17.12, p < .001) [inconsistency], but not the Position factor, their interaction, or any of the interactions involving the Condition factor (p > .277 in all of the cases).

N400 effects

A Linear Mixed Model was performed on the Pz electrode using the time window of 300 to 500 ms from the onset of the final word (Mestres-Missé et al., 2007), with Condition as the factor (5 levels, HC congruent, HC incongruent, LC congruent, LC incongruent and NS). The predicted main effect of Condition was significant (F(4, 73.82)= 20.75, p < .001). Once verified that mean voltages changed significantly as a function of the condition, we further investigated the differences carrying out a new repeated-measures Linear Mixed Model including Cloze-Probability (2, High and Low) and Congruency (2, Congruent and Incongruent) as factors.

The analysis revealed significant main effects of Cloze-Probability (F(1,84) = 20. 12, p < .001) and Congruency (F(1,61.99)= 68.44, p < .001). The Cloze-Probability × Congruency interaction was also significant (F(1,58.58) = 8.00, p = .006). Post-hoc tests revealed that incongruent endings to the HC condition produced larger amplitudes (M = –3.55) than congruent endings (M = –0.44) (p <.001, 95% CI of the difference [2.29 3.92]), just as the amplitude of incongruent endings to the LC condition (M = –4.5) was larger than that of congruent endings (M = –3.13) (p = .001, 95% CI of the difference [0.55 2.19]). Within congruent endings, the mean voltage of HC congruent was significantly more negative than LC congruent (p < .001, 95% CI of the difference [1.67 3.70]), while LC incongruent was not significantly different from the

HC incongruent (p = .064, 95% CI of the difference [-0.06 1.97]). Regarding the NS control condition (M = -2.45), its mean amplitude was significantly more negative than HC congruent (p = .008, 95% CI of the difference [0.45 2.86]), but it was not significantly different from LC congruent (p = .06, 95% CI of the difference [-0.04 2.11]). The incongruent minus congruent difference waveforms were obtained for both Cloze-Probability conditions on the Pz electrode. Onset latency was determined and subjected to statistical analyses as described in the main text. The difference wave for HC endings diverged from 257 to 829 ms (for a 572 ms duration, $2.29 < t(21) < 7.37$) and the difference wave for LC endings diverged from 338 to 824 ms (a duration of 486 ms, $2.11 < t(21) < 8.24$). A comparison between the difference waveforms of the two conditions revealed significant differences from 304 to 465 ms (a total duration of 161 ms, $2.18 < t(21) < 2.86$). The topographical mapping of the temporal evolution of the congruency effect also evidenced the latency difference (see Fig. S4) and revealed a typical centro-parietal distribution of the N400 (Kutas & Hillyard, 1984).

Study 2

Ahead of time: early sentence slow cortical modulations associated to semantic prediction

This study corresponds to:

León-Cabrera, P, Flores, A., Rodríguez-Fornells, A-, & Morís, J. (2019) Ahead of time: early sentence slow cortical modulations associated to semantic prediction. *NeuroImage*, 189, 192-201. DOI: https://doi.org/10.1016/j.neuroimage.2019.01.005

6 Study 2

Ahead of time: early sentence slow cortical modulations associated to semantic prediction

6.1 Introduction

Sentences unfold linearly in time. As every word is encountered, it is incorporated into a broader representation to achieve comprehension rapidly (van Berkum, Brown & Hagoort, 1999; van Petten, Coulson, Rubin, Plante & Parks, 1999), in line with our everyday experience of understanding language on the fly. This immediacy might be bolstered by predictive mechanisms, whereby top-down information guides the pre-activation of upcoming words to facilitate their processing upon receipt (for a review, Kutas, DeLong & Smith, 2011). In the past years, extensive research has shown that words that are more predictable given a previous context show facilitated processing over low predictable ones, as indexed by reduced amplitudes in the N400 component (e.g. van Petten et al., 1999; Federmeier et al., 2007; Kutas & Federmeier; 2011), and there is evidence that, under some circumstances, readers may pre-activate specific words (Wicha, Bates, Moreno & Kutas, 2003; Delong, Urbach & Kutas, 2005; van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005; Freunberger & Roehm, 2016; but see Nieuwland et al., 2018). Conversely, other authors have argued that, given the infinite number of potential continuations to the same phrase, prediction is an unviable strategy in language and that comprehension proceeds in a strictly bottom-up, stimulus-driven fashion (e.g., Jackendoff, 2003, 2007; Morris, 2006). More recently, some researchers have called into question whether prediction is actually necessary for language comprehension (see Huettig & Mani, 2016). To contribute in the quest for uncovering whether and when predictive processing may occur in language, the present study investigated electrophysiological correlates associated to the build-up of predictions over the course of sentence comprehension.

In order to investigate the prediction process *per se*, one strategy is to focus on the anticipatory stage of the words that are being predicted. To do so, sentences are built so that their semantic content establishes a weak or a strong expectation for the final word. Highly constraining sentences (HC) have a single best completion (e.g. “the dentist proceeded to clean her… teeth”), whereas low constraining (LC) sentences have more than one likely continuation (e.g. “the meeting was arranged for the… morning/afternoon/evening). Importantly, several studies have shown that the brain activity preceding highly predictable endings (i.e. “teeth”) is different from that prior to less predictable ones (i.e. “morning”), supporting that the information from the context changes the state of the cognitive system before critical words are encountered. In a recent ERP study in speech comprehension, we reported the development of a slow negative potential in the anticipatory period prior to closing words, with larger amplitudes for HC than for LC and non-semantic sentences (León-Cabrera, Rodríguez-Fornells & Morís, 2017). Interestingly, this neural signature was reminiscent of the stimulus-preceding negativity (SPN), a slow potential that has been described in other domains as an index of anticipation for upcoming relevant events (for reviews, see Brunia, van Boxtel & Böcker., 2011; Hackley, Valle-Inclấn, Masaki & Hebert, 2014). More specifically, it showed a similar gradual amplitude increase over time, a graded modulation that agrees with the predictability of the upcoming stimulus and a frontal topographical distribution. Given its characteristics and the task that was employed, we suggested that it might be a language-related SPN reflecting the semantic anticipation of the upcoming word. In line with this, Grisoni, Miller and Pulvermüller (2017) reported a slow negative potential prior to the final words for HC (but not LC) sentences. They interpreted it as an index of specific semantic predictions, given that it emerged over the motor areas specifically associated to the upcoming concept –dorsolateral for hand-related verbs (e.g. “write”) and ventral for face-related verbs (e.g. “talk”). Despite the task differences, both studies converge in finding anticipatory slow potentials that are modulated by semantic constraint in a way that is consistent with contextually-driven anticipation.

The investigation of brain activity in the anticipatory period of words has also proven fruitful using different methodological approaches (Dikker & Pylkkänen, 2013; Piai, Roelofs & Maris, 2014; Rommers, Dickson, Norton, Wlotko & Federmeier, 2017; Wang, Hagoort & Jensen, 2018). However, exploring the anticipatory period of one word offers a very restricted time interval that can only reveal rather transient processes. According to interactive models of language comprehension, contextual information has an immediate and continuous impact on bottom-up processing (e.g. Nieuwland & van Berkum, 2006). Therefore, the effect of semantic constraint should be detectable much earlier in time and be captured by a sustained, rather than punctate brain processes. Hereafter, we will use the labels “local” and “sentence-level” to distinguish between these times.

Interestingly, ERPs are well-suited to track the time-course of sentence-level dynamics. In the past, several ERPs studies have found cumulative, sentence-level slow modulations that are distinguishable from faster, word-related processes (for a review, Kutas, 1997). Most of the available multiword ERP studies focused on sentences of varying syntactic complexity. In these, slow sustained positivities were found for sentences with simpler structures which are easy to integrate (Kutas & King, 1996), whereas negativities were associated to increased working memory (WM) demands when processing more complex sentences (in reading comprehension, King & Kutas, 1995; Fiebach, Schleswesky & Friederici, 2002; in speech comprehension, Müller, King & Kutas, 1996). The involvement of WM was further backed-up by a correlation between the amplitude of the slow sustained negativity and individual differences in WM capacity – participants with a lower WM span showed larger negativities than those with a higher span (see also, Vos, Gunter, Schriefers & Friederici, 2001). On the other hand, other WM operations can also result in sustained negativities. As such, Münte, Schiltz & Kutas (1998) found that the amount of conceptual information activated by a single word led to a similar modulation.

With regards to the current goal, a previous study compared sentence-level dynamics between different levels of semantic constraint, and did not find differences (Kutas, van Petten & Besson, 1988). Critically, however, the experiment was originally designed to investigate the N400 component (see, Kutas & Hillyard, 1980) and thus there are several methodological issues (e.g. averaging sentences with different word lengths) that might have precluded the observation of significant slow wave effects. Using an appropriate task design, we hypothesized that the effect of semantic constraint would be captured by slow sustained modulations, arising through sentence comprehension. Furthermore, we expected to find the same pattern of amplitude differences as that of local slow negative potentials associated to semantic anticipation during spoken sentences (León-Cabrera et al., 2017). Such a correlate would be consistent with prediction-based models of sentence comprehension, whereby the probabilities of upcoming continuations are computed and updated gradually as the semantic context accrues (for a review, Kutas et al., 2011).

The goal of this study was to investigate sustained ERP correlates associated to the build-up of predictions over the course of sentence comprehension. To address this, we adapted the task of the previous study using spoken sentences (León-Cabrera et al., 2017) for visual presentation, which also allowed us to answer additional questions. First, we presented sentences using rapid visual serial presentation (RVSP) to equate the duration of words and track word-by-word changes by temporally aligning them. Critically, this feature of the design allowed a simple and straightforward analysis of the EEG activity during sentence processing. This analysis was not possible in the previous study, given that the stimuli were presented auditorily and, consequently, the sentences had variable durations. Second, the use of a different input modality (i.e. visual) allowed us to test the generality of the findings in the previous study. That is, whether similar results regarding the local negativity and the N400 component would replicate in the visual modality.

All the other features of the task and the materials employed were the same as in León-Cabrera et al. (2017). The sentences had different levels of semantic constraint (HC, LC or none) and the final word appeared after a delay of 1000 ms, in order to extend its anticipatory period. We inserted the delay in the attempt to (1) replicate previous findings of local slow potentials time-locked to the onset of the anticipatory period, and, in case we found a sentence-level modulation, (2) determine whether both levels of analysis (i.e. sentence-level and local) captured the same cognitive operations. Based on our previous study, we expected more negative amplitudes for HC than LC contexts. Finally, to assess the potential consequences of prediction and misprediction, the final word turned out to be incongruent in half of the trials so as to evoke the N400 component. If some representation of the final words become pre-activated at some point, an earlier onset of the N400 effect in the HC condition could be observed (León-Cabrera et al., 2017), as an index of processing facilitation after supportive contexts.

6.2 Methods

Participants

Twenty-four right-handed young adults (12 females, M = 23.2, SD = 4.3) were paid to participate in the experiment after giving written consent. We discarded the data of three participants due to excessive blinking (n = 1) and signs of sleepiness as evidenced by high alpha activity throughout the experiment (n = 2). Therefore, the total sample included 21 participants (11 females, M = 23.4, SD = 4.5). None of the participants reported health problems or prior neurological disorders.

Stimuli and Task

The three-hundred fifty-two context sentences (176 high constraint and 176 low constraint) used in León-Cabrera et al. (2017) and created by Mestres-Missé, Rodríguez-Fornells & Münte (2007) were employed (**Table 1**). According to the original paper, the mean cloze-probability sentence completions was 6.1% (SD = 10.3) for the low constraint condition (LC) and 76% (17.7%) for the high constraint condition (HC). To construct a non-semantic condition (NS), we randomly picked 40 sentences from the LC condition and scattered the vowels within each word to make them semantically meaningless but grammatically plausible. The non-semantic condition was included to provide a control condition in which the sentence was formed by linguistic stimuli of similar complexity to the words of the HC and LC conditions, but without any semantic content. Each context sentence had one congruent final word (their best completion) and one incongruent and implausible final word. Given the features of the congruent final words (mean word length = 6.6 letters [SD = 1.87], mean number of syllables = 2.9 [SD = 0.82] and mean word frequency = 35.18

[SD = 75.18]), we carefully selected the incongruent final words from the ESPAL database (Duchon, Perea, Sebastián-Gallés, Martí & Carreiras, 2013) so that they matched the congruent words (mean word length = 6.6 letters [SD = 1.87], mean number of syllables = 2.88 [SD = 0.88] and mean word frequency = 33.68 [SD = 69.34]). Also matching the features of the congruent words, another intact 40 final words were chosen for the non-semantic condition (mean word length = 6.72 letters [SD = 1.77], mean number of syllables = 2.93 [SD = 0.76] and mean word frequency = 33.42 [SD = 61.74]).

Table 1. Sample set of sentence examples for each experimental condition with translations to English.

Condition	Sentence Context	Final word
HCC	El portero fue capaz de atrapar la The goalkepper managed to catch the	pelota ball
HCI	El portero fue capaz de atrapar la The goalkepper managed to catch the	orilla shore
LCC	Le ha regalado a su hijo una As a present she gave her son a	pelota ball
LCI	Le ha regalado a su hijo una As a present she gave her son a	orilla shore
NS	Helade algoroa seujohi nua	trip (viaje)

We controlled potential confounds of the effect of contextual constraint in the N400 component. To do that, final words always had the same congruency status (congruent or incongruent), but whether they followed HC or LC contexts was counterbalanced across participants. For instance, half of the participants read "plane" as a congruent continuation for the HC context "I have never flown on a…" and the other half as a congruent ending for the LC context "The dot in the sky must be a…". Likewise, the word "drink" served as an incongruent word for the same sentences. Furthermore, every congruent and incongruent pair (i.e. "plane" and "drink") were matched in word length, word frequency, familiarity, imaginability and concreteness. In the case of the NS condition, the scattered sentence ended with a different word (e.g. "trip") that was matched in the same variables with the corresponding congruent and incongruent words (e.g. "plane" and "drink") in the semantic conditions.

Trials proceeded as follows (**Figure 1**). First, a fixation point (a cross) appeared at the center of the screen for a period of between 1350 and 1750 ms (uniform distribution with a 50 ms step).

Then the sentence appeared, presented one word at a time for 200 ms (500 ms stimulus onset asynchrony, SOA). The font used was Courier New, with a size of 36 points. The color of the letters was black, while the background was white. A 1 s delay took place after the offset of the penultimate word, to analyze the anticipatory period of the final word. After the end of the final word, they waited for 800 ms until the blinking signal appeared (a depiction of an eye at the center of the screen). The blinking signal remained for 2 s before the next trial began.

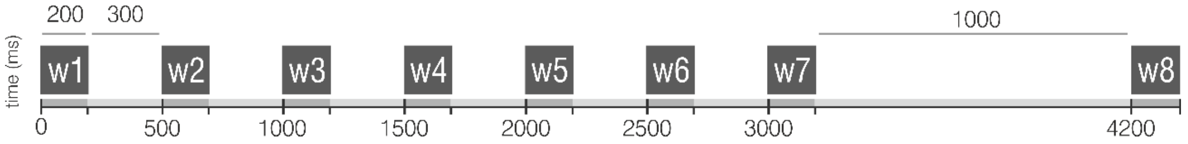

Figure 1. Depiction of the structure of a trial. Sentences were presented one word at a time on a screen. Each word was displayed for 200 ms. A 300 ms SOA was inserted between the words in the sentence context (from w1 to w7) and a 1000 ms SOA (i.e. pre-word interval) separated the penultimate (w7) and the final word (w8).

To ensure that participants were reading attentively, a memory recognition test followed each block. In this test, ten words were visually presented one at a time; half were old words and half were new ones. Participants had to respond, by pressing a key, whether they had read that word in the previous block or not. Every word remained on the screen until an answer was supplied, followed by a 600 ms fixation cross before the presentation of the next word. The correspondence between the keys ("z" and "m") and the responses ("yes" and "no") were counterbalanced across participants. At the end of the memory task, the main task continued right away, except for even-numbered blocks, which were followed by a pause that could be resumed anytime.

Procedure

After general instructions and preparation, participants were comfortably seated in approximately 70 cm away from the computer screen that would be later used. After this, the EEG cap was set up and the state of each electrode was checked. Participants were given instructions about how to reduce artifacts by minimizing movement and to wait for a visual signal at the end of each trial to blink. After completing another unrelated task, the briefing was carried out. Participants were told to read each sentence carefully and that after each block they would have to complete a recognition test related to the final words presented during those trials.

EEG Recording

Electrophysiological data (EEG; sampling rate = 500 Hz; on-line bandpass filter = .015 – 1000 Hz) was recorded from 29 tin scalp electrodes at standard 10/20 system positions (Jasper, 1958;

electrode positions: FPz, FP1/2, Fz, F3/4, F7/8, FCz, FC3/4, Cz, C3/4, CPz, CP3/4, Pz, P3/4, TP7/8, T3/4, T5/6, Oz, O1/2, left and right mastoids). The EEG signal was re-referenced offline to the mean activity of the mastoid electrodes. Vertical and horizontal ocular electrodes were recorded and used for artifact rejection. All electrode impedances were kept below 5 kΩ. Before performing statistical analysis, the data were filtered offline at 50 Hz and 60 Hz with a notch filter (to attenuate electrical line noise) and at 30Hz using a low-pass Butterworth filter (roll-off of 12 dB/oct) as implemented in ERPLAB toolbox V6.1.3 (López-Calderón & Luck, 2014). To perform artifact rejection, we excluded the epochs in which the peak-to-peak amplitude in the ocular electrodes exceeded ± 85 µV (moving window = 200 ms, moving step = 20 ms) or in which activity was ± 200 µV in any other channel. After this, additional visual inspection of the resulting signal was carried out for each subject individually. Using these criteria, a mean of 19.05% of trials were rejected (SD = 11.7).

ERP Data Analysis

The data analyses were divided into three parts that focused separately on the electrophysiological activity, 1) over the course of the sentence context (sentence-level interval), 2) in the anticipatory period of the final word (pre-word interval), and 3) at final word processing (post-word interval). We applied repeated measures ANOVA to perform confirmatory analysis on the effects for which we had a prior hypothesis based on previous studies, and cluster-based permutation to carry out exploratory analysis and/or to adequately control for Type I errors when multiple comparisons were involved.

Sentence-level interval. Sentence processing was examined within an epoch length of 4200 ms that comprised the activity from the onset of the first word (w1) to the end of the delay period before the eighth and final word (w8). The epoch was baseline-corrected to its preceding 100 ms. For this analysis, we had a hypothesis regarding the direction of the differences between conditions based on a previous study (León-Cabrera et. al, 2017) that showed more negative amplitudes to HC. However, as we had no strong prediction about the spatiotemporal locus of the effect in the time-course of the whole sentence, we adopted an exploratory approach and applied a non-parametrical cluster-based permutation test (Maris & Oostenveld, 2007) to analyze experimental differences in this period. This method controls the probability of a false positive (Type I error rate) in the set of multiple comparisons. Following this procedure, every sample (Channel x Time) was compared between two conditions (HC vs. LC, HC vs. NS and LC vs. NS, separately) by means of *t*-statistics. Then, an algorithm clustered the adjacent spatio-temporal samples with similar differences based on a *t*-value threshold of ±1.72 (alpha level of 0.05 with 20 degrees of freedom, for one-tailed testing). A cluster-level statistic (permutation *p*-value) was computed under a permutation distribution of the cluster with the largest sum of t-values. The permutation distribution was approximated by a Monte Carlo method involving 15000 randomizations of the data of between the two experimental conditions. Only clusters with a permutation *p*-value below 5% (critical alpha level) were considered significant. We chose one-tailed testing because –as previously stated– we had an a priori hypothesis that the amplitude would be more negative for HC compared to all the other experimental conditions (HC < LC < NS). Before running the test, we applied a low-pass filter at 5Hz to focus solely on slow brain potentials (Brunia et al., 2011). Importantly, the cluster-based permutation procedure deals effectively with the increased probability of noise autocorrelation involved in setting such a small low-pass filter cut-off (Piai, Dahlslätt, & Maris, 2015) and therefore it is a very adequate statistical solution for these situations.

Pre-word interval. Based on previous studies (Morís, Luque & Rodríguez-Fornells, 2013; León-Cabrera et al., 2017), we expected differences between the experimental conditions to be maximal at frontocentral sensors in the period immediately preceding the final word. We analyzed the 1000 ms delay period preceding the final word (w8) in a separate window analysis, time-locked to the onset of the penultimate word (w7) and using the 100 ms pre-stimulus as a baseline (Van Petten & Kutas, 1991; King & Kutas, 1995). The differences were assessed by pre-planned repeated measures analysis of variance (ANOVA) of three time-windows in the final 600 ms of the delay period (–600 to –400 ms, –400 to –200 ms, and –200 to 0 ms) (Morís et al., 2013) on a subset of six frontocentral sites.

Post-word interval. The analysis of the activity elicited by the presentation of the final word focused on the N400 component. The time window for the analysis comprised the 300 to 500 ms time-locked to the onset of the final word (w8), with a 100 ms pre-stimulus baseline (León-Cabrera et al., 2017). The CPz electrode data was analyzed using a repeated-measures ANOVA. CPz was chosen as the representative electrode based on the topographical distribution of the strongest N400 effect in the sample, although, for completeness, an analysis using a cluster of centroparietal electrodes (Cz, CPz and Pz) was also carried out (Supplementary Materials). We expected to find an earlier onset of the N400 congruency effect for HC compared to LC (León-Cabrera et al., 2017). Given that we had an a priori hypothesis about the location of the effect, we run a single-sensor cluster-based permutation test (Maris & Oostenweld, 2007; Bullmore et al., 1999) on the CPz electrode to determine the temporal onset of the effect for level of semantic constraint (HCI minus HCC, and LCI minus LCC, separately) in the 650 ms after final word onset. Again, we chose one-tailed testing because we predicted that incongruent words would have more negative mean amplitudes than congruent words. We did not plan to include the NS condition in these analyses because its functional interpretation is unclear in the post-word interval, we nevertheless added it in a grand-averaged ERP in the Supplementary Materials for those interested.

Statistical analyses were run with the FieldTrip toolbox (version 28-01-2018) (Oostenveld, Fries, Maris & Schoffelen, 2011) and with SPSS 21. In repeated-measures ANOVA, the Greenhouse-Geisser correction was applied whenever the sphericity assumption was not met. Corrected *p*-values are reported as p_{GG}. In case a theoretically relevant interaction was found to be significant, we disentangled the effect by means of post-hoc *t* tests.

6.3 Results

Sentence-level slow sustained negativity

The grand average ERPs of the sentence context (**Fig. 2**) revealed widespread slow brain potentials with amplitude differences across experimental conditions that were confirmed statistically by the non-parametric cluster-based permutation test (**Fig. 3**).

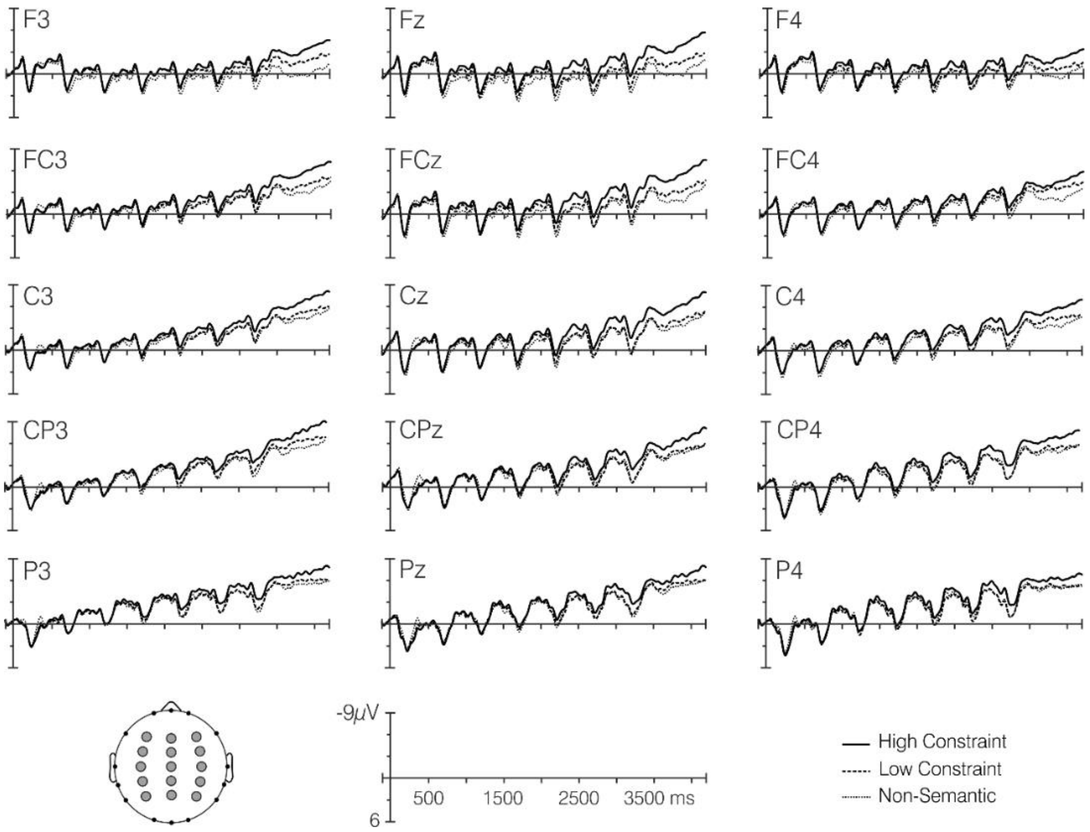

Figure 2. Grand-averaged ERPs across sentence contexts (0 to 3200 ms) and the following final pre-word anticipatory period (3200 to 4200 ms) showing the sentence-level slow cortical modulations at frontal (F3, Fz, F4), frontocentral (FC3, FCz, FC4), central (C3, Cz, C4), centroparietal (CP3, CPz, CP4) and parietal (P3, Pz, P4) sites in the three experimental conditions (High Constraint, HC; Low Constraint, LC and Non-Semantic, NS). Negative is plotted upward.

A significant negative cluster was found both between HC and LC (p <.001) and HC and NS (p <.001), consistent with the hypothesis that voltage amplitudes would be more negative for higher levels of semantic constraint. In the HC and LC contrast, the cluster started around the fifth word (w5; from 2200 ms onwards), whereas it started earlier in the HC and NS comparison, around the fourth word (w4; from 1600 ms onwards). In both cases, the cluster prolonged to the end of the sentential context and appeared to be maximal at frontocentral sites. No significant clusters were found between LC and NS. As a measure of the robustness of these effects, we replicated the finding of these significant negative clusters using a more conservative testing (Supplementary Materials).

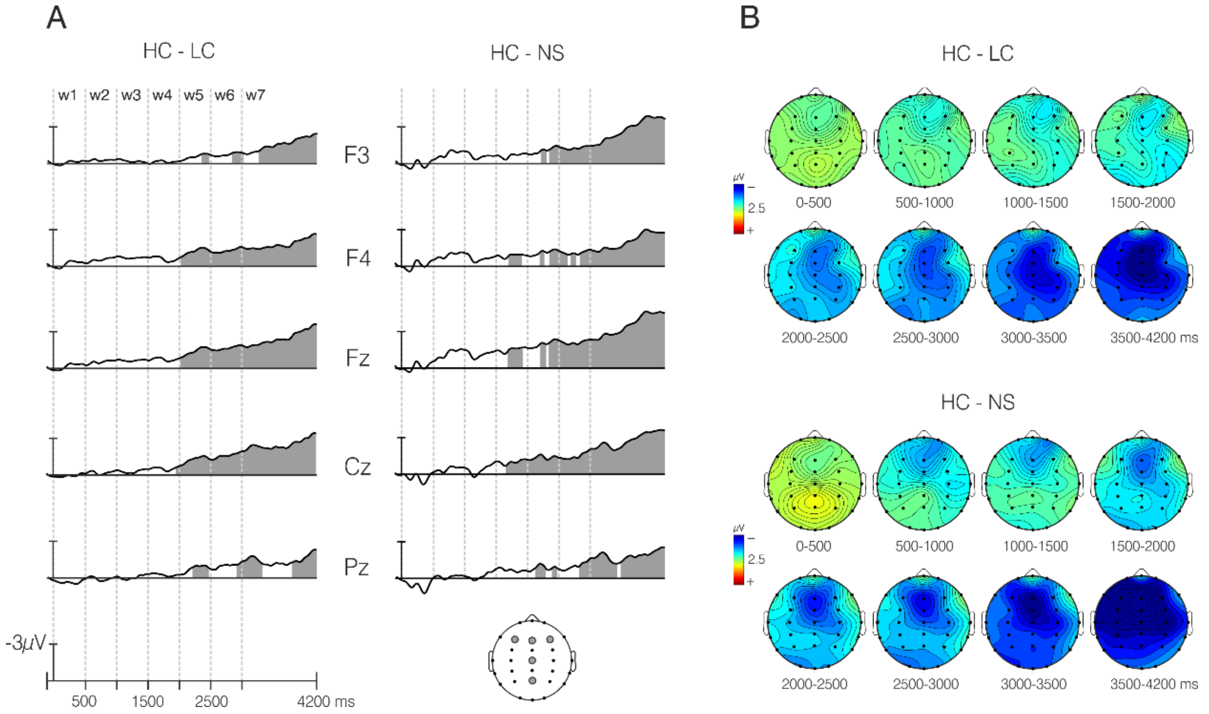

Figure 3. A: Difference waveforms of the statistically significant contrasts (HC minus LC and HC minus NS) in the non-parametrical cluster-based permutation analysis that comprised the full epoch (0 to 4200 ms) time-locked to the onset of the sentence context, at 5 representative electrodes (F3, F4, Fz, Cz and Pz). The onset of each word is marked with a vertical dotted line and the corresponding word position (w_x). The grey portions represent the statistically significant areas of the cluster. Negative is plotted upward. **B:** Scalp maps displaying the temporal evolution of the sentence-level sustained negativity based on mean amplitude differences (HC minus LC and HC minus NS) in eight consecutive epochs of 500 ms. The range of voltage values for the maps is ± 2.5 μV.

Local slow negative potential

We investigated the delay preceding the presentation of the final word, as we expected brain activity to portray the maximal differences in the slow preceding negativity during this period (**Fig. 4**).

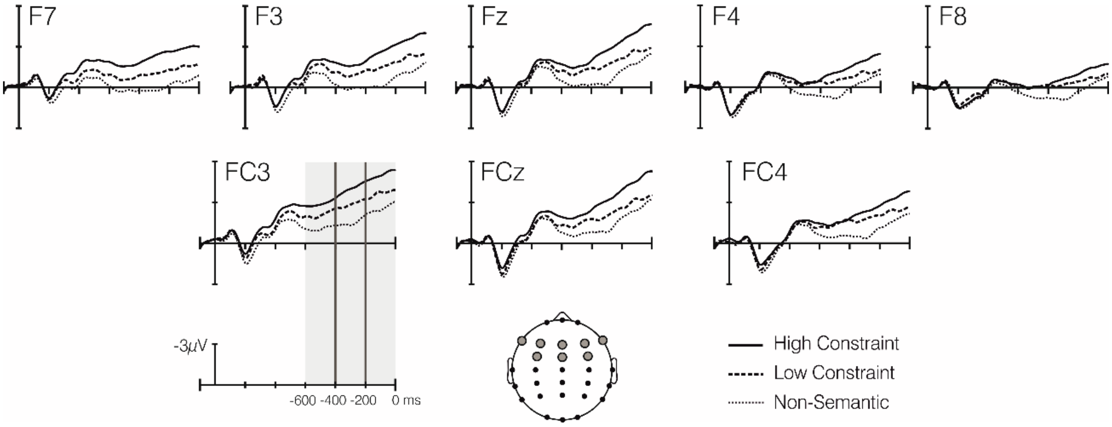

Figure 4. Grand-averaged ERPs to the pre-word interval showing the local slow negative potential preceding the presentation of the final word at frontal (Fz, F3/4, F7/8) and fronto-central (FCz, FC3/4) sites, where the differences were maximal. The ERPs are time-locked to the onset of the penultimate word. The grey area indicates the time intervals that were separately subjected to statistical analysis. The time values (X axis) are referenced to the onset of the final word.

A repeated-measures ANOVA was conducted on the mean amplitudes of six fronto-central electrodes (F3, Fz, F4, FC3, FCz and FC4), involving four factors: Condition (3 levels, HC, LC, NS), Time (3 levels, 400 to 600 ms, 600 to 800 ms, and 800 to 1000 ms), Electrode (2 levels, frontal, frontocentral), and Laterality (3 levels, left, central, right). Given the computation of multiple comparisons, Bonferroni-corrected post-hoc *t* tests were carried out to further explore the effects found. The corrected *p*-values are indexed as p_{BF}.

Overall mean amplitude differences between each level of semantic constraint were quantified as a main effect of Condition ($F(2,40) = 19.08$, $p < .001$, $n^2_p = 0.48$). Equally to what we observed in the whole sentence analysis, the mean amplitude linearly changed with the level of semantic constraint ($F(1,20) = 70.75$, $p < .001$, $n^2_p = 0.78$), with HC being significantly more negative than LC ($p_{BF} = .031$) and NS ($p_{BF} < .001$), and LC being comparatively more negative than NS ($p_{BF} = .036$).

Consistent with the typical evolution of the slow stimulus preceding negativity, there was a significant main effect of Time ($F(2,40) = 17.05$, $p_{GG} < .001$, $n^2_p = .46$) that stemmed from a steeper negative shift in the mean amplitude of the final 200 ms, as can be seen in the scalp maps (**Fig. 5**). More specifically, *t* tests showed that the mean amplitudes between the first and second time-windows did not differ, but the third time-window (corresponding to the last 200 ms) became significantly more negative (p_{BF} <.001 in both contrasts). Indeed, the Condition x Time interaction was significant ($F(4,80) = 3.09$, p =.02, $n^{2}_{p} = .13$). In the earliest time-window, from –600 to –400 ms, HC and LC did not differ, whereas both were more negative than NS (HC vs. LC, p = .242, HC vs. NS, p < .001, LC vs. NS, p = .036). Later, from –400 to –200 ms, all the conditions differed significantly between them. Again, as observed in the main effect of condition, showing more negative amplitudes in HC compared to LC (p_{BF} =.009) and NS (p_{BF} <.001), and LC being more negative than NS (p_{BF} =.004). Eventually, in the final 200 ms, HC remained more negative than LC (p_{BF} =.002) and NS (p_{BF} <.001), but the difference between LC and NS disappeared.

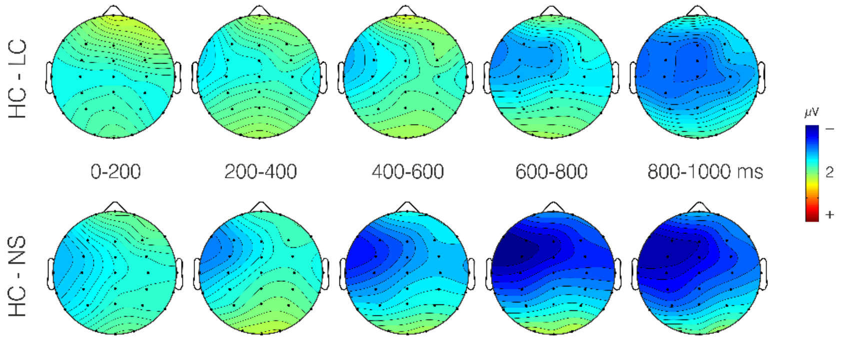

Figure 5. Scalp maps of the difference waveforms (HC minus LC and HC minus NS) showing the temporal evolution of the local slow negative potential that developed in the 1000 ms pre-word interval in five consecutive time-windows of 200 ms. The range of voltage values for the maps is ± 2 μV.

Finally, the significant Condition x Laterality interaction ($F(4,80) = 2.74$, p = .03, $n^{2}_{p} = .12$) confirmed the prediction that the effect of semantic constraint was most pronounced at left sites. In fact, at right and central positions only HC and NS diverged (p_{BF} =.02), whereas all the conditions were different at left sites (all p_{BF} <.05). There were also main effects of Electrode (p_{BF} <.001) and Laterality (p_{BF} =.002), and significant interactions of Position x Laterality, Time x Electrode, Time x Laterality, and Time x Position x Laterality (all p_{BF} <.05). We do not expand on these results because they do not interact with the effect of the experimental manipulation and thus are considered uninformative in this context.

N400 component

We conducted a repeated-measures ANOVA on the CPz electrode data using the time window of 300–500 ms from the onset of the final word to investigate the N400 component (**Fig. 6A**), using two factors, contextual constraint (2 levels, high, low) and congruency (2 levels, congruent, incongruent). The analysis showed significant main effects of constraint, ($F(1,20) = 30.9$, p <.001, $n^{2}_{p} = .60$), and congruency ($F(1,20) = 30.32$, p <.001, $n^{2}_{p} = .60$) , but no significant interaction between the factors. Post-hoc t tests revealed the classical N400 effect, whereby incongruent words produced more negative amplitudes than congruent words, both within high (mean amplitudes, HCI = 2.65, HCC = 6.53, $p < .001$) and low (mean amplitudes, LCI = –.16, LCC = 2.19, $p = .002$) constraint conditions. To test the robustness of the effect, we computed an equivalent ANOVA using a centroparietal cluster (Cz, CPz and Pz) that replicated the same results (Supplementary Materials).

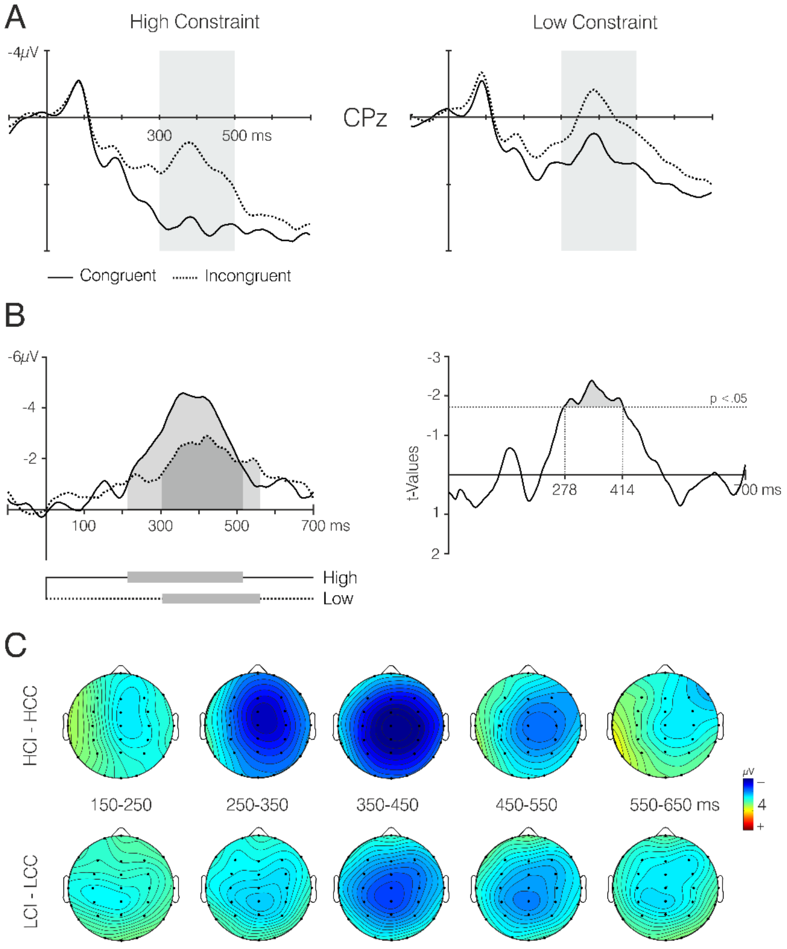

Figure 6. A: Grand-averaged ERPs for each condition showing the N400 component at the CPz electrode, time-locked to the onset of the final word. The grey area indicates the interval where the amplitudes were statistically compared. **B. Left**: Difference waveforms (Incongruent minus Congruent) displaying the N400 effect to each condition of semantic constraint. The grey segments correspond to the statistically significant clusters. **B. Right**: Difference waveforms of t-value evolution of the differences between the two waveforms. The grey dotted line indicates the point at which the waves start to differ. C. Scalp maps of the mean amplitude differences every 100 ms in the 150 to 650 ms interval, showing the earlier onset of the effect for the HC compared to the LC condition. The voltage values for the maps ranged ± 4µV.

Single-sensor cluster-based permutation tests on the difference waveforms of the constraint conditions (HCI minus HCC, and LCI minus LCC) at CPz, confirmed the prediction that congruency effects started earlier for high than low levels of semantic constraint (**Fig. 6B. Left**). A significant cluster was found from 216 to 522 ms in the high constraint contrast and from 302 to 564 ms in the low constraint condition (all cluster *p*-values < .001, one-tailed testing). Lastly, we report the time segment in which *t* values of the comparison between the two difference waveforms exceed the critical value under a one-tailed *t*-distribution (±1.72, for one-tailed testing). The period comprises a window from 278 to 414 ms (136 ms duration) (**Fig. 6B. Right**). The topographical representation of the difference waveforms visibly exhibits the distinct temporal evolution of the effect (**Fig. 6C**).

6.4 Discussion

In this study, we investigated sustained ERP correlates associated to the build-up of predictions during sentence processing. We found a sentence-level slow and sustained negativity that developed gradually over the course of sentence processing, capturing the differences in semantic constraint. This sustained modulation showed more negative amplitudes for higher than lower levels of semantic constraint. In the pre-word interval, the anticipatory period elicited a local slow negative potential with a left-lateralized distribution immediately prior to the final word. Once the final word was presented, we found a canonical N400 modulation to semantic fit and an earlier onset of the N400 effect for high compared to low contextual constraint.

The pattern of results in the pre-word interval strongly resembled the findings in a previous study using a similar task in speech comprehension (León-Cabrera et al. 2017). In both studies, we observed the development of a left-dominant slow negative potential over frontocentral sites, with a graded pattern of amplitude differences as a function of contextual constraint, more negative amplitudes for stronger levels of constraint (HC < LC < NS). In that study, we noted that the observed potential shared many features with the SPN – a frontal locus, a gradual amplitude increase over time, and a modulation that agrees with the predictability of the upcoming event (Walter et al., 1964; Brunia and Damen, 1988). Following this, we interpreted the slow potential as an index of semantic anticipation associated to the upcoming word. Now, the fact that we found a very similar slow potential in two different input modalities (i.e. auditory and visual) bolsters the idea that the underlying mechanism of this brain signal might be associated to high-level operations, such as semantic processes. If it reflected generic anticipatory attention to the word based on its temporal predictability, we would have expected to find different topographical distributions for each sensory modality (Brunia & Van Boxtel, 2004).

Interestingly, the topographical distribution of the local slow potential is strikingly similar to those of previous reports of slow potentials linked to the retrieval of verbal representations. Heil, Rösler & Henninghausen (1996) had participants learn the associations between a drawing and one, two or three words. During recall, a frontal negativity developed prior to the presentation of the word. Resembling the current findings, the slow wave was maximal at left anterior areas (i.e. F3 electrode), where the amplitude was monotonically related to the number of verbal representations that had to be retrieved (see also Rösler, Heil & Henninghausen, 1995). Other studies have associated left anterior negativities to the elaborative encoding of verbal associations (Lang et al., 1988), the retention of verbal material (Ruchkin, Johnson, Canoune & Ritter, 1990) or the active maintenance of verbal items (Khader, Ranganath, Seemüller & Rösler, 2007). More generally, in studies of language comprehension using more elaborate materials, transient left anterior negativities have been typically associated to WM operations such as the storage or retrieval of specific elements (Kluender and Kutas, 1993; Fiebach, Schlesewsky & Friederici, 2002; Piai, Meyer, Schreuder & Bastiaansen, 2013; Matzke, Mai, Nager, Rüsseler & Münte, 2002). Considering this, the local negativity could be associated to the retrieval and/or the maintenance of a verbal representation, which would fit with a stronger form of prediction – once the potential candidates have been narrowed down, if the context is sufficiently constraining, the most probable continuation might be accessed to facilitate its imminent bottom-up processing (for a review, Kutas et al. 2011). However, the current task does not allow us to know exactly what information (if any) might be activated. Of note, Grisoni et al. (2017) attributed to semantic prediction their finding of a slow negative potential that was maximal over the motor areas conceptually related to the final word. The distinct topographical distribution may be accounted by several task differences, such as the use of action verbs (instead of nouns, in our case) or the restriction of candidates to only two possible semantic categories (instead of an unbounded set of categories). It is possible that the different task dispositions led to the activation of qualitatively distinct information, thus tapping on distinct cortical regions.

Importantly, the present study extends the previous findings by revealing that the local slow potential may be the continuation of a slower, sentence-level process. Based on interactive models of language comprehension, we hypothesized that contextual constraint would have an earlier and sustained impact on brain activity during sentence processing. Accordingly, we found that HC contexts diverged already at the fourth word from NS contexts, and at the fifth word from LC contexts. These differences were captured by slow sustained negativities that developed gradually across sentence processing and that were broadly distributed over frontal and central areas. On the one hand, the sustained negativity could be a result of the sum of successive word-level integration processes, given that it spanned over several words. Under this view, it might reflect the incremental construction of a meaning representation. Accordingly, the pattern of differences could be explained by the relative ease or difficulty integrating the words that each type of context affords; the greater semantic information provided by HC contexts, relative to LC and NS contexts, would have facilitated the incorporation of the words into the higher-level representation of the sentence, and possibly also the achievement of a clearer message-level interpretation. On the other hand, the sentence-level modulation might be capturing more than word-by-word context updating. The differences are also consistent with more recent prediction-based models of language comprehension, whereby contextual information influences the state of the language processing system before bottom-up input is received (for a review, Kutas et al., 2011). In fact, the early and progressive evolution of the differences agrees well with the notion that predictions about upcoming concepts (and/or some of their features) are computed and continuously updated online (for a review, Kuperberg & Jaeger, 2016).

As we described in the Introduction, high-level integration during sentence processing has been associated with slow positive (rather than negative) shifts (see Kutas & King, 1996, for a review). For instance, van Petten & Kutas (1991) found that reading semantically meaningful and congruent sentences lead to a word-by-word decrement of the N400 component, resulting in a slow positive shift. Cross-clause negativities have been linked to increased WM demands when processing sentences that have more difficult syntactic configurations (e.g., King & Kutas, 1995; Müller et al., 1996; Fiebach et al., 2002) or in which the linguistic order of presentation of the events did not match the conceptual order (Münte et al., 1998). Given the differences between the tasks and materials employed, it is difficult to assess to what extent the same cognitive mechanisms might have played a role in this study. But based on the previous ERP reports of sentence processing dynamics, the polarity of the sentence-level modulation might be indicative of differences in memory demands as a function of contextual constraint. For instance, some authors have suggested that memory demands could be involved in maintaining multiple representations active in parallel when the context allows for several interpretations (Wlotko & Federmeier, 2016). Also, during predictive processing, memory demands could be involved in maintaining predicted representations available within the cognitive system, at least until the true bottom-up input is found. In this line, some accounts of prediction have suggested that items are predictively added to the contextual representation that is being held in WM (Lau, Holcomb & Kuperberg, 2013) and therefore imply a potential interaction between working memory and predictive processes.

In the sentence interval, the LC and the NS condition did not differ significantly. This may seem surprising at first, given that it was impossible for participants to build an internal meaning representation in the NS condition, whereas they should have been able to do so in the LC condition. However, it should be noted that the LC condition in this task had a very low constraint

– the mean cloze probability was below 10%. Therefore, it is possible that participants could not accrue much information online and were unable to build a robust context representation in the LC condition. In a similar vein, if we interpret the effects in terms of prediction, participants might not have found useful semantic cues to generate predictions online. At this point, we should also consider that some features of the task – such as the cloze-like structure of the sentences (Ferreira and Lowder, 2016), word-by-word presentation of stimuli or/and the introduction of a delay – may have encouraged participants to adopt a more strategic approach for sentence comprehension that would not be engaged in other situations (Brothers, Swaab & Traxler, 2017). For example, if the task structure had been less favorable for prediction, or the comprehension had assisted an ulterior task goal, perhaps the qualitative distinction between NS and LC sentences would have been more critical and the two conditions would have differed substantially during sentence processing.

Finally, we examined processing differences once the final word was presented and found the classical effect of N400 amplitudes graded by word expectancy (Kutas and Hillyard, 1984). In addition, the N400 effect (that is, unexpected minus expected word) had an earlier onset in the HC than in the LC condition. Importantly, this replicates a previous report of the same effect using auditory stimuli (León-Cabrera et al., 2017), arguing strongly against previous claims stating that the N400 latency is not affected by psychological factors (Federmeier & Laszlo, 2009). Based on our experimental manipulation, the latency effect could stem from facilitated processing for congruent words in the HC contexts (Federmeier & Kutas, 1999). In HC contexts, the sentence context would bias the language system strongly towards the best upcoming continuation, and thus, if the bottom-up input eventually matches the prediction (i.e. congruent ending), word integration may proceed faster and yield an earlier N400 component. On the other hand, under stronger views of prediction, HC contexts would lead to the pre-activation of the lexical representation of the best completion (Boudewyn, Long & Swaab, 2015). The availability of these features would allow detecting a mismatch sooner when the input is incongruent and consequently elicit an earlier N400 response in HC than in LC contexts. Although the present data is insufficient to tell apart the exact underlying mechanism, the finding of a latency effect strengthens the involvement of semantic prediction during comprehension.

Altogether, these findings contribute significantly to the available literature by providing clear evidence of early, continuous and gradual impact of top-down information on sentence comprehension. The sentence-level negativity represents one of the few instances of slow and sustained changes associated to semantic constraint, extending previous studies that focused on more transient effects (León-Cabrera et al., 2017; Grisoni et al., 2017). Overall, the pattern of results adds to previous evidence in support of prediction in language comprehension, in the sense that the state of the system changes prior to bottom-up input (in this case, at least prior to the final word of the sentence) as a function of top-down information. Interestingly, the local slow potential could be reflecting retrieval or maintenance of some verbal representation, although we did not find clear evidence of predictive pre-activation of a specific lexical representation neither in the pre- nor at the post-word intervals. Future research is needed to investigate the potential relationship between the observed slow negativities and stronger forms of prediction, as well as differences in memory demands. Insofar, the current experimental approach seems to be a valuable device to reveal potential markers of predictive processes engaged during sentence comprehension, it remains an open question whether these results can be transferred to different experimental setups or to more naturalistic comprehension contexts. Insofar, the current experimental approach seems to be a valuable device to reveal potential markers of predictive processes engaged during language comprehension, and future studies should address whether these results can be transferred to different experimental set-ups or to more naturalistic comprehension contexts.

6.5 Supplementary materials

Two-tailed cluster-based permutation analysis

Every sample (Channel x Time) was compared between two conditions (HC vs. LC, HC vs. NS and LC vs. NS, separately) by means of t-statistics. The adjacent spatio-temporal samples were clustered based on a t-value threshold of ±2.08 (alpha level of 0.05 with 20 degrees of freedom, for two-tailed testing). A cluster-level statistic (permutation p-value) was computed under a permutation distribution of the cluster with the largest sum of t-values. The permutation distribution was approximated by the Monte Carlo method involving 15000 randomizations of the data of between the two experimental conditions. Only clusters with a permutation p-value below 5% (critical alpha level) were considered significant.

A significant negative cluster was detected between HC and LC ($p<.001$) as well as between HC and NS ($p<.01$). The differences started earlier in the HC vs. NS contrast (fourth word) than in the HC vs. LC contrast (fifth word), and prolonged until the end of the sentence in both conditions. No significant clusters were found in the comparison between the LC and the NS conditions. These results replicate the findings using one-tailed testing that are reported on the main paper.

N400 effect using a centroparietal cluster (Cz, CPz and Pz)

We conducted a repeated-measures ANOVA on a centroparietal cluster (Cz, CPz and Pz) using the time window of 300–500 ms from the onset of the final word with two factors: contextual constraint (2 levels, high, low) and congruency (2 levels, congruent, incongruent) as factors. The analysis yielded significant main effects of constraint, ($F(1,20) = 33.2$, $p <.001$), and congruency ($F(1,20) = 33.4$, $p <.001$) , but no significant interaction between the factors ($F(1,20) =3.64$, $p = .07$). Post-hoc t-tests revealed that incongruent words produced more negative amplitudes than congruent words, both within high ($p < .001$) and low ($p = .001$) constraint conditions.

Study 3

Alpha oscillatory activity associated with predictive language processing in written and spoken comprehension

7 Study 3

Alpha oscillatory activity associated with predictive language processing in written and spoken comprehension

7.1 Introduction

Language prediction is thought to have an important role in achieving fast and efficient comprehension (Kuperberg, Brothers & Wlotko, 2020; Kuperberg & Jaeger, 2016; Kutas & Federmeier, 2011). This is grounded in extensive evidence that people predict different aspects of incoming input during comprehension, including semantic (Federmeier & Kutas, 1999), grammatical (Wicha, Moreno & Kutas, 2004; van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005) or phonological features (DeLong, Urbach & Kutas, 2005; but see, Nieuwland et al., 2018). Yet, the neural mechanisms that support and enable language prediction are still unclear.

Neural oscillations – patterns of synchronization and desynchronization of neural activity captured in the EEG and MEG signal as power increases and decreases – have been linked to predictive language processing (Lewis and Bastiaansen, 2015). A common strategy to study prediction in language comprehension is to manipulate the amount of semantic information of sentence contexts, such that they are strongly predictive of one specific word, or not (high constraint contexts, HC or low constraint contexts, LC; for examples, see Table 1). Using this approach, oscillations in theta (4-7 Hz) and gamma (>40 Hz) bands have been attributed to the match and mismatch of lexical and semantic predictions of critical words (Wang, Zhu & Bastiaansen, 2012; Hald, Bastiaansen and Hagoort, 2005; Bastiaansen & Hagoort, 2015). By contrast, prediction implies that processing differs before the critical input. Accordingly, recent studies in written comprehension have observed alpha (8-12 Hz) and/or beta (~13-20 Hz) power decreases prior to critical words in HC relative to LC, that is, when contexts afford strong predictions about the upcoming word (Rommers, Dickson, Norton, Wlotko & Federmeier, 2017; Li, Zhang, Xia & Swaab, 2017; Wang, Hagoort & Jensen, 2018; Terporten, Schoffelen, Dai, Hagoort & Kösem, 2019). However, the exact nature of these pre-word alpha-beta power decreases remains unclear, and the findings diverge across studies (Molinaro, Monsalve & Lizarazu, 2016; Prystauka & Lewis, 2019).

A question that remains unanswered is to what extent the aforementioned pre-word alpha and beta power decreases are dependent or independent of the modality of comprehension, which can be informative about their underlying mechanisms. On the one hand, pre-stimulus alpha and beta power decreases have been associated with processes that are dependent on the sensory modality, such as sensory gating (Jensen, Bonnefond, & VanRullen, 2012) and sensory prediction of upcoming input (Arnal & Giraud, 2012). Consistent with these views, alpha power decreases over posterior regions involved in processing an imperative visual stimulus (Thut et al., 2006; Worden, Foxe, Wang, & Simpson, 2000; Foxe et al., 1998) or over temporal sites when the imperative stimulus is auditory (for a review, see Weisz, Hartmann, Müller, Lorenz & Obleser, 2011). On the other hand, besides sensory processing, recent accounts propose that pre-stimulus alpha-beta power decreases reflect modality-independent operations that support information processing as well (Hanslmayr, Staudilg, & Fellner, 2012; van Ede, 2018; Griffiths et al., 2019). Specifically, alpha-beta power decreases are proposed to provide an optimal brain state for information processing (Hanslmayr, 2012), which can account for their correlations with successful memory encoding (Klimesch, 1997; Khader & Rösler, 2011; Hanslmayr et al., 2011). Also, alpha and beta power decreases are associated with lexico-semantic retrieval before naming a picture that is preceded by an HC context in both visual and auditory modalities (for a review, see Piai & Zheng, 2019).

Given this background, we examined to what extent alpha and beta power decreases during language prediction in sentence comprehension reflect modality-dependent and modality-independent processes. To test this, we characterized oscillatory activity in the pre-existing datasets of two EEG studies in spoken (León-Cabrera, Rodríguez-Fornells & Morís, 2017) and in reading comprehension (León-Cabrera, Flores, Rodríguez-Fornells & Morís, 2019). Oscillatory activity was computed in the one-second interval prior to sentence-final words in HC or LC contexts (see **Table 1**) in each study individually. Then, both studies were qualitatively and quantitatively compared. Based on previous findings, we expected to find alpha and/or beta decreases prior to words in HC relative to LC contexts.

7.2 Methods

Reading comprehension task

Participants. Twenty-four young adults (12 females) voluntarily participated in the experiment in exchange for monetary compensation. The data of two participants was removed due to excessive blinking and movement artefacts, resulting in a final dataset of 22 participants (12 females; M = 23.3 years [SD = 4.4]). All participants were native Spanish speakers, right-handed, with normal or corrected-to-normal vision and no history of neurological problems. The study was approved by the Ethics Committee of Hospital Universitari de Bellvitge.

Materials. The materials were the same in both studies. A total of 176 high constraint (HC) and 176 low constraint (LC) sentence contexts were employed (created and categorized by Mestres-Missé, Rodríguez-Fornells and Münte, 2007). The sentence contexts were originally designed to have the same final word (best completion) in both conditions. For example, the LC sentence "The dot in the sky must be a plane" shared the final word with the HC sentence "I have never flown on a plane". In other words, the two sets were matched in their final words. For every participant, a set of 80 HC sentences and 80 LC sentences were pseudo-randomly selected from the list with the condition that none shared the final word. As a result, each final word was encountered only once during the experiment while their constraint status was counterbalanced across participants. Half of the sentences in each constraint condition ended with a congruent word (the best completion) and half with an incongruent and implausible word. The mean cloze probability of the final words was 6.1% (SD ± 10.3%) in the LC condition and 76% (SD ± 17.7%) in the HC condition (**Table 1**). Incongruent endings were selected from the ESPAL database (Duchon et al., 2013) so that they matched the congruent endings in mean word length, mean number of syllables, word frequency, familiarity, imaginability and concreteness. Finally, a non-semantic condition (NS) served as a control condition in which no meaning could be derived from the sentence contexts (further specifications about this condition can be found in the original papers). We note, however, that this condition will not be analysed in the present study.

Table 1. Example sentences of each condition (original sentences in Spanish in brackets).

Condition	*Sentence context*	*Congruent*	*Incongruent*
HC	The goalkeeper managed to catch the (El portero fue capaz de atrapar la)	ball (pelota)	shore (orilla)
LC	As a present she gave her son a (Le ha regalado a su hijo una)	ball (pelota)	shore (orilla)

Procedure. Participants were comfortably seated approximately 70 cm away from the computer screen. The EEG cap was set up and the state of each electrode was checked. Before the experiment started, participants were briefed on the importance of minimizing movement and were instructed to synchronize their blinks to the blinking signal and to prevent them at other times. Participants were told to attentively read the sentences and that they would have to complete a short recognition test of the final words at the end of every block of sentences. Following the instructions, the experimental task started (see **Figure 1** for a schematic display of the trial structure).

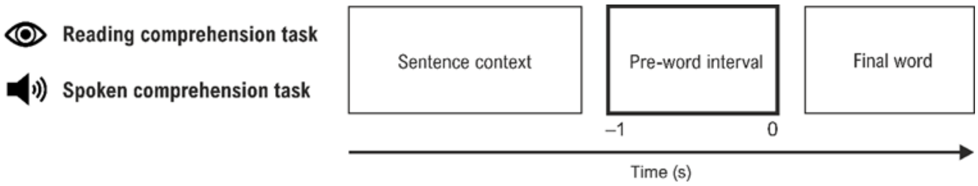

Figure 1. Simplified trial structure highlighting (outlined in black) the 1-second long pre-word interval in which the analyses will be carried out. The sentence contexts (high or low constraining) and final words (congruent or incongruent) were presented visually (one word at a time) or auditorily in the Reading and Spoken Comprehension task, respectively.

The experiment consisted of 10 blocks of 20 trials (4 HC congruent, 4 HC incongruent, 4 LC congruent, 4 LC incongruent and 4 NS). The order of the sentences within each block was random. Words were black on white background (font type: Courier New; font size: 36 points). On each trial, a fixation point (a cross) appeared at the centre of the screen for a variable interval (between 1.35 and 1.75 s in a uniform distribution with a 50 ms step). Then, a sentence was presented word-by-word (200 ms per word, followed by a 300 ms blank screen). Between the penultimate and the final word, the blank screen remained for a period of 1 s (pre-word interval). After the final word (200 ms duration), a blank screen was presented (800 ms) before the blinking signal (2 s; depiction of an eye at the centre of the screen). The next trial began after the offset of the blinking signal.

To ensure that participants were reading attentively, a recognition test followed each block. In this test, ten words were visually presented one at a time. Half were final words that they had encountered in the previous block (randomly selected) and half were unseen words. They pressed a keyboard key (“z” or “m”) to indicate whether they had read that word in the previous block or not. Each word remained on the screen until an answer was supplied, followed by a 600 ms fixation cross before the presentation of the next word. The correspondence between the key (“z” and “m”) and response type (“yes” and “no”) was counterbalanced across participants. The next block started immediately after the test except for even-numbered blocks, which were followed by a resting period that could be resumed anytime by the participant.

Data acquisition. Electrophysiological data (EEG; sampling rate = 500 Hz; on-line bandpass filter = .015–1000 Hz) was recorded from 31 tin scalp electrodes at standard 10/20 system positions (electrode positions: FPz, FP1/2, Fz, F3/4, F7/8, FCz, FC3/4, Cz, C3/4, CPz, CP3/4, Pz, P3/4, TP7/8, T3/4, T5/6, Oz, O1/2, left and right mastoids). Vertical and horizontal electro-oculograms (outer canthus and infraorbital ridge of the right eye) were recorded to monitor eye blinks and movement. All electrode impedances were kept below 5 kΩ.

EEG pre-processing. The same pre-processing procedure was applied in both studies. The EEG analyses were performed using FieldTrip version 20181231 (Oostenweld, Fries, Maris, & Schoffelen, 2011) in MATLAB R2019b. First, data were notch-filtered at 50 Hz to attenuate electrical line noise and re-referenced offline to the average of the mastoid electrodes. The continuous data were segmented into epochs encompassing the activity from –1000 ms to 2000 ms time-locked to the pre-word interval onset (i.e. 1000 ms before and after the pre-word interval). After this, Independent Component Analysis was applied to remove eye movement artefacts (as implemented in FieldTrip). Between 1 and 3 components were removed. Additionally, any remaining epochs with artefacts (eye blinks, eye movements, electrode drifting or muscle activity) were rejected through visual inspection before trials were split by condition.

After artefact rejection, the percentage of valid trials in the Reading Comprehension task was 90.2 % in the HC condition (mean = 72 trials; std = 6.28 trials), and 89.65 % in the LC condition (mean = 71 trials; std = 5.26 trials).

Time-resolved power analyses. Time-frequency (TF) representations of power were computed on the pre-word interval segment (a 3 s segment ranging from – 1 s pre-word to 2 s post-word) encompassing the frequencies from 2 to 30 Hz. A moving window short-time Fast Fourier Transform (FFT) approach was used. A sliding window of three cycles was advanced in steps of 20 ms and 1 Hz. The data in each time window was multiplied with a Hanning taper and Fourier transformed. The obtained spectrograms were averaged over trials for each participant and condition. For contrasts between conditions, each condition was normalized by the average across conditions (i.e. for every participant, the power spectrum of each condition was divided, element by element, by the averaged power spectrum of the two conditions) (for other studies using the same approach, see Piai, Roelofs & Maris, 2014; Rommers et al., 2017; Rommers & Federmeier, 2018). We used this approach because it allows to assess relative power changes preventing pre-stimulus baseline activity from influencing power estimates, which was appropriate in this case given that differences between the conditions were expected to be present before the pre-word interval due to the processing of the penultimat word and prior sentence context (León-Cabrera et al., 2019). For visual inspection of power changes in individual conditions in the pre-word interval, we computed the power change relative to a baseline from −500 ms to −150 ms (for prior studies using the same parameters, see Wang et al., 2012; Rommers et al., 2017).

We tested for effects of contextual constraint in the whole pre-word interval of 1 s using a non-parametric cluster-based permutation test (Maris & Oostenveld, 2007) as implemented in FieldTrip. All frequencies, time points and channels were blindly scanned for adjacent frequencies, time points and channels with a similar difference across conditions. Importantly, this method successfully controls the probability of a false positive (Type I error rate) in the set of multiple comparisons. Following this statistical procedure, every sample (time x frequency x channel) was compared between the two conditions (HC vs. LC) by means of a paired-samples t-statistic. Channels were set to have an average of 6.1 neighbours. Then, adjacent time-frequency-channel samples were clustered based on a t-value threshold of ± 2.07 (i.e., an alpha level of 0.05 with 21 degrees of freedom). The permutation p-value was computed using the Monte Carlo method involving 5000 random permutations. Only clusters with a permutation p-value below 5% (two-tailed testing) were considered significant.

Spoken comprehension task

Participants. Twenty-two young adults (12 females; M = 21.1 years [SD = 1.8]) voluntarily participated in the experiment in exchange of monetary compensation. All participants had normal hearing, were native Spanish speakers, right-handed, with normal or corrected-to-normal vision and no history of neurological problems. The study was approved by the Ethics Committee of Hospital Universitari de Bellvitge.

Materials. The materials were the same as the in the Reading comprehension task (see Section 0). For auditory presentation, all the linguistic stimuli were transformed into audio using a voice-synthesizer software (Loquendo TTS Director, 2005). This software creates natural-sounding audio, very close to natural speech, while allowing precise control of the speech rate, the amplitude and the prosody of the speech (for a sample sentence of each condition, see the supplementary materials of León-Cabrera et al., 2017). After audio conversion, sentence contexts had a mean duration of 1966 ms (SD = 288 ms) (HC =2000 ms, LC =1925 ms and NS =1992 ms). Final words had a mean duration of 490 ms (SD =122 ms) (Congruent words = 485 ms; Incongruent words = 494 ms; Neutral words = 495 ms).

Procedure. The procedure followed was the same as in the Reading Comprehension task, but the task was adapted for auditory presentation. Trials proceeded as follows (see **Figure 1** for a schematic depiction of the trial structure). A fixation point (a black cross on white background) appeared and remained at the centre of the screen throughout the trial. Participants were instructed to fixate their gaze on it. Two seconds after the presentation of the fixation point, the sentence was auditorily presented with a 1 s silent interval between the penultimate and the final words (pre-word interval). One second after the offset of the final word the blinking signal was presented (a depiction of an eye at the centre of the screen). The blinking signal remained for 2 s before the next trial began.

Data acquisition. The only differences compared to the description of the Reading Comprehension task were the sampling rate (sampling rate = 1000 Hz) and the montage, which included 31 tin scalp electrodes at standard 10/20 system positions as well, but with a different configuration (electrode positions: Fp1/2, Fz, F3/4, F7/8, FCz, FC1/2, FC5/6, Cz, C3/4, T3/4, T5/6, CP1/2, CP5/6, Pz, P3/4, PO1/2, Oz, left and right mastoids).

EEG pre-processing. The pre-processing used was the same as in the the Reading Comprehension task. After artefact rejection, a total of 91.3 % valid trials were retained in the HC condition (mean = 73 trials; std = 6.55 trials) and of 93 % in the LC condition (mean = 74 trials; std = 5.43 trials).

Time-resolved power analyses. The analysis pipeline was the same as in the Reading Comprehension task.

7.3 Results

To visualize the power changes for the individual conditions (see **Figure 2**), we computed power changes relative to the baseline from –500 to –150 ms time-locked to the onset of the pre-word interval (see Wang et al., 2012; Rommers et al., 2017).

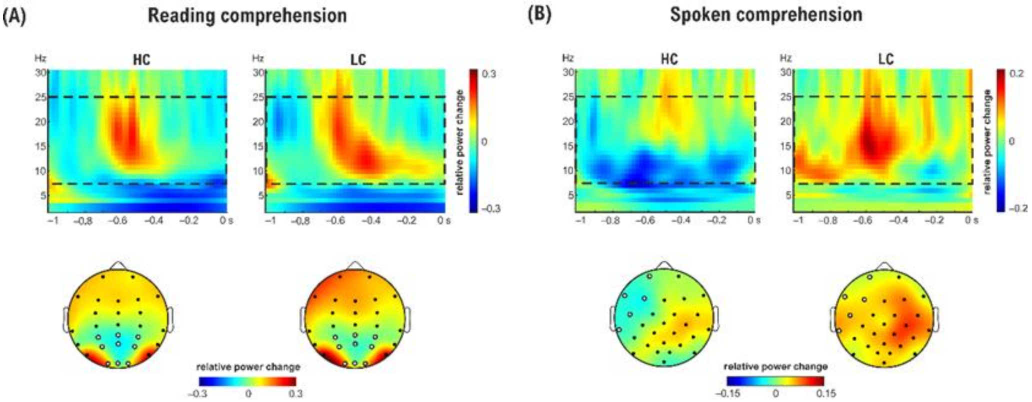

Figure 2. Grand-averaged TFRs and topographical maps of the individual conditions for the two tasks. The colour scale indicates the proportion of power change relative to a pre-stimulus baseline of –500 to –150 ms. The final word appeared at 0 sec. **A)** Reading Comprehension. **Top**: TFRs of the individual conditions at the subset of channels highlighted in the scalp maps (CP3/4, CPz, P3/4, Pz, O1/2, Oz). The discontinuous line indicates the alpha-beta range (8-25 Hz). **Bottom**: corresponding scalp maps of averaged power in the alpha-beta band (8-25 Hz). **B)** Spoken comprehension. **Top**: TFRs of the individual conditions at a subset of representative channels highlighted in the scalp maps below (FP1, F3, F7, C3, T3). **Bottom**: corresponding scalp maps of the averaged power in the alpha-beta band (8-25 Hz).

Because the baseline contains activity of the processing of the penultimate word, no baseline normalization was performed for contrasts between conditions (read Section 2.1.6 for details). Therefore, the spectrograms of the contrasts (**Figure 3** and **Figure 4**) do not derive directly from those of the individual conditions (**Figure 2**).

Reading comprehension task

Visual inspection of the relative power changes as a function of sentential constraint revealed a transient decrease in power in the alpha band for HC relative to LC (**Figure 3**). The relative difference in power was supported by the finding of a significant cluster (p = 0.0378), detected in the 8-12 Hz alpha band. This result is displayed in **Figure 3B**, which shows the relative power decrease for the HC relative to the LC condition (top) and the statistically significant cluster displayed as highlighted activity (bottom) averaged over all the channels associated with the significant cluster. Although the scalp distribution of the cluster was widespread, the effect was most prominent over frontal and parieto-occipital sites (see **Figure 3B**). Further exploration of the timing of the cluster shown in **Figure 5**. A. revealed that the effect had a mainly frontocentral distribution from –600 to –400 ms, and later, from –400 to –200 ms, was more constricted to parieto-occipital sites.

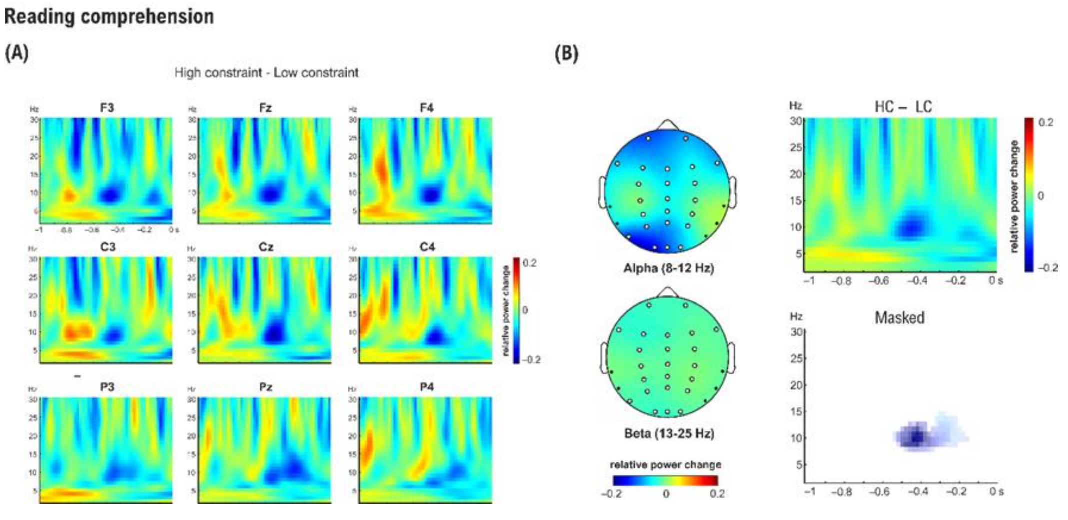

Figure 3. Grand-averaged TFRs and topographical maps of the sentential-constraint effect (i.e. high minus low) in the Reading Comprehension task. The final word appeared at 0 sec. **A)** TFRs of the effect of sentential constraint (high minus low constraint, normalised by the average of the conditions) at representative set of frontal, central and parietal channels. **B)** Scalp maps of the averaged power in the pre-word interval in the alpha (8-12 Hz) and beta (13-25 Hz) bands. Highlighted channels (white dots) are those that showed differences in the statistical analysis. **Top**: TFR map of the effect of sentential constraint at all channels associated with the significant cluster (highlighted in the scalp maps). **Bottom**: thresholded TFR showing the significant spectrotemporal cluster for the sentential constraint effect.

Spoken comprehension task

As can be seen in **Figure 4A**, the spectrograms of the relative power changes between the conditions suggested a similar power decrease for HC (relative to LC) in the upper alpha and lower beta bands (from 8 to 20 Hz approximately), particularly over left anterior sites, but these apparent differences were not confirmed by the planned cluster-based permutation test.

Following this observation, at the expense of increasing the probability of a Type I error, we decided to perform an ad-hoc analysis in which we increased the sensitivity of the test. We considered it was justified to further explore the apparent alpha power decrease given its convergence with previous reports (Rommers et al., 2017; Wang et al., 2018), as well as with the present findings for the Reading Comprehension task. In addition, there are several reasons to expect less robust effects in the auditory domain (which we elaborate on in more detail in the Discussion). The goal of this ad-hoc test was to confirm that, if there is any effect, it is in the alpha band, as the visual inspection suggests. To do this, we parcelled the scalp into similar areas (left fronto-central set: FP1, F7, F3, Fz, FC1, FC5, Cz, C3, T3; right fronto-central set: FP2, F8, F4, Fz, FC2, FC6, Cz, C4, T4; left centro-posterior set: CP1, CP5, Pz, P3, PO1, Oz; right centro-posterior set: CP2, CP6, Pz, P4, PO2, Oz) and performed a cluster-based permutation test on each set of electrodes. In all tests, all frequencies and time points were included, and all other parameters were the same as in the planned tests. With this procedure, a statistically significant cluster was detected in the left fronto-central set ($p = 0.031$), capturing a power decrease in the HC (relative to LC) condition, encompassing the upper alpha and lower beta bands (approximately from 10 to 15 Hz) and most prominent from –700 to –500 ms before the presentation of the final word (see **Figure 4B**. for the scalp map of the effect in this time window). The additional tests, including all frequencies and time points, were carried out on electrode subsets of similar size covering the remaining brain regions (as explained above). No significant clusters were found scanning any of these subsets (all p-values $>.05$).

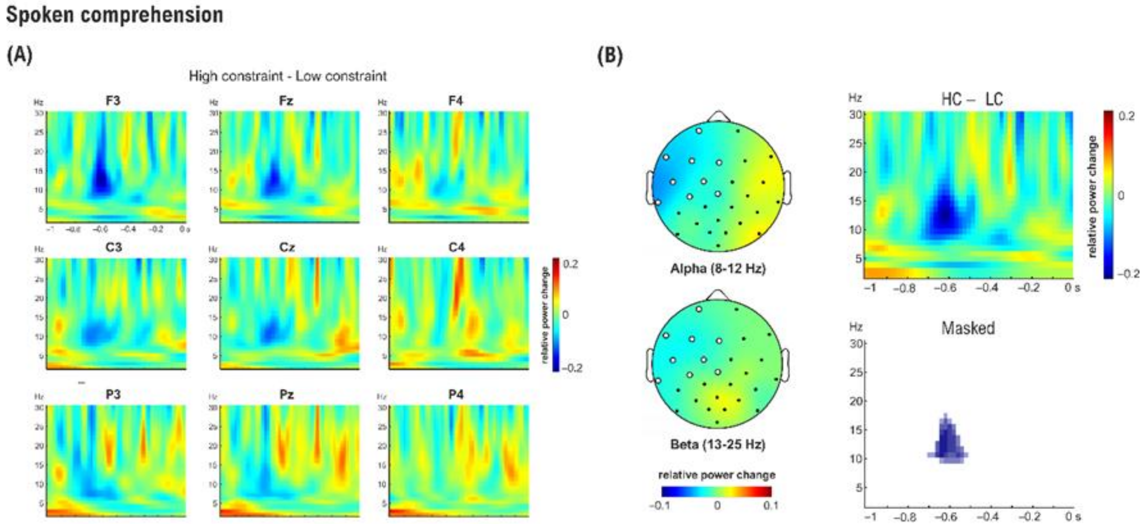

Figure 3. Grand-averaged TFRs and topographical maps of the sentential-constraint effect (i.e. high minus low) in the Spoken Comprehension task. The final word appeared at 0 sec. **A)** TFRs of the effect of sentential constraint (high minus low constraint, normalised by the average of the conditions) at representative set of frontal, central and parietal channels. **B)** Scalp maps of the averaged power in the pre-word interval in the alpha (8-12 Hz) and beta (13-25 Hz) bands. Channels that showed significant differences, although less robust, in the ad-hoc statistical analysis are highlighted (white dots). **Top**: TFR of the effect of sentential constraint at all channels that showed a significant cluster (highlighted in the scalp maps). **Bottom**: Thresholded TFR showing the significant spectro-temporal cluster for the effect of sentential constraint.

Between-experiment comparison

The main results in the two tasks are displayed together in **Figure 5** for visual comparison. Overall, significant alpha power decreases preceded the presentation of predictable lexical sentence endings (relative to unpredictable endings) in both tasks but with some differences with respect to their spectro-temporal dynamics, topographical scalp distribution and robustness. Perhaps the most prominent distinction is the topographical distribution of the alpha power decrease, with parieto-occipital channels involved only in the Reading Comprehension modality (**Figure 5A**). Regarding spectral dynamics, the power decrease extended beyond the hypothesized alpha band (8-12 Hz) onto the lower beta band (13 to 20 Hz) in the Spoken Comprehension task, whereas it remained more constrained to the alpha band range in the Reading Comprehension task. Lastly, the significant cluster capturing the effect was detected earlier and was more short-lived in the Spoken than in the Reading Comprehension task.

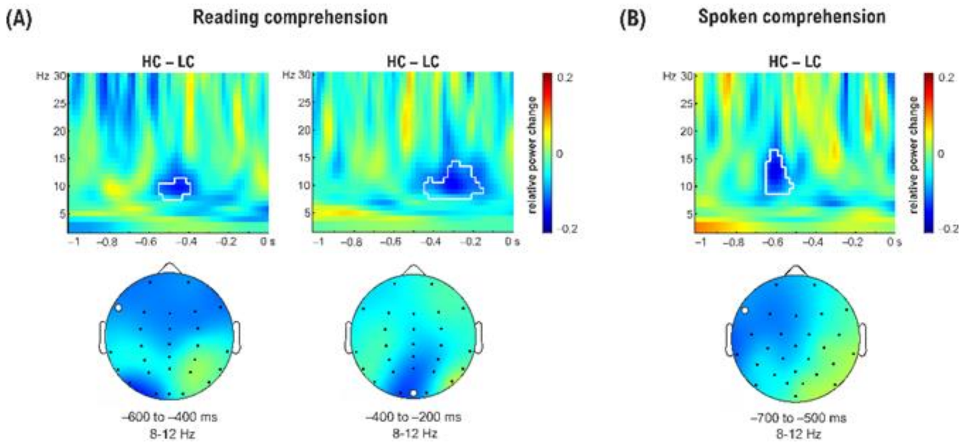

Figure 5. Grand-averaged TFRs and topographical maps of the sentential-constraint effect at single representative channels (marked in white) in both tasks for qualitative comparison. The significant clusters are indicated as outlined spectro-temporal regions (in white). Power decreases for the HC (relative to the LC) condition in the alpha (and sometimes lower beta) band are evident in both tasks but with different timing and topographical distribution. **A) Top.** TFRs of the effect in the Reading Comprehension task at F7 (right) and Oz (left). **Bottom.** corresponding scalp maps of alpha (8-12 Hz) modulations averaged over the temporal windows in which the effect was most consistent in each electrode. **B) Top.** TFRs of the effect in the Spoken Comprehension task at F7. **Bottom.** corresponding scalp map in the alpha-band (8-12 Hz) and in the time interval were the effect was maximal.

Altogether, the qualitative comparison of the effects reveals a similarity between the two experiments, namely, a frontally distributed alpha power decrease, but also differences with regards to the timing and scalp distribution. To obtain further information about the extent to which the effects may be similar, we used a between-subjects cluster-based permutation test to statistically compare the effects (i.e. normalized difference HC minus LC in the pre-word interval) between the two experiments. Given the different electrode montage used in each experiment, this analysis was performed only on the shared electrodes (18 electrodes: FP1/2, F3/4, C3/4, P3/4, F7/8, T3/4/5/6, Cz, Fz, Pz, Oz). The analysis revealed no significant clusters between the two experiments, indicating that there is a certain degree of similarity between the oscillatory activity of the two experiments.

We additionally inspected the group-level power spectra as a function of contextual constraint over separate subsets of electrodes (left frontocentral, right frontocentral, centroparietal and parieto-occipital) in each task (**Figure 6**). We focused on the alpha band where maximal effects were observed in the present study. This additional analysis might be helpful to illustrate the robustness of the findings in the time-frequency domain by providing convergent information using a procedure that does not require data normalization. As can be seen in the figure below, the delta band (below 4 Hz) contain the largest proportion of spectral power, in line with the 1/f typical pattern of brain activity. Besides that, spectral power was condensed on the upper theta (6-7 Hz) and especially the lower alpha band (8-10 Hz). The differences between the conditions follow the same pattern as in the time-frequency analyses, namely, decreased power when contextual constraint was higher (HC) relative to when it was lower (LC). The scalp distribution of the differences in absolute power also looks consistent with the effects found in the time-frequency domain. In the Reading Comprehension task, differences are widely spread with a seemingly greater divergence over occipital electrodes, whereas in the Spoken comprehension task, the relative decrease of HC is most prominent in the subset comprising left anterior electrodes.

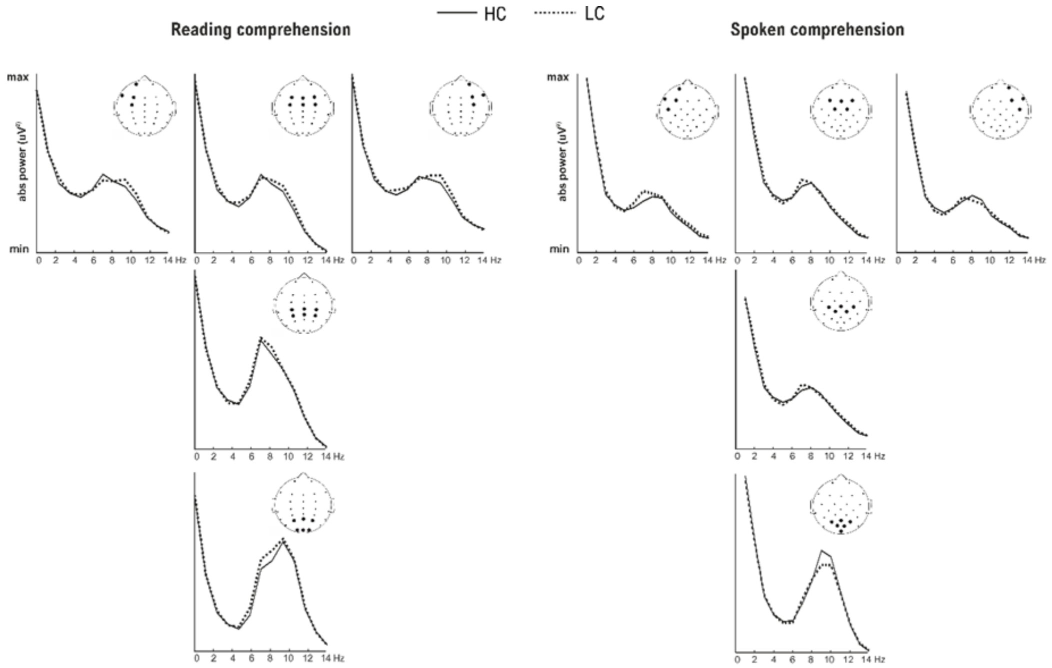

Figure 6. Time-averaged power spectra as a function of sentential constraint for the two tasks Reading Comprehension (left) and Spoken Comprehension (right) tasks, averaged over five sets of electrodes highlighted next to each spectrum. The power axis (y-axis) is scaled to span from 0 to maximum power for each spectrum individually.

7.4 Discussion

The present study examined to what extent alpha and beta power decreases during language prediction in sentence comprehension reflect modality-dependent and modality-independent processes. To test this, a conjoint analysis of two studies using the same materials in different modalities of comprehension (i.e., reading and spoken comprehension) was conducted. Participants read or listened to HC and LC sentence contexts that afforded strong or weak predictions about a sentence-final word and that could match or mismatch the contextual prediction. In reading comprehension, alpha (8-12 Hz) decreased in HC relative to LC sentences in the interval prior to the critical word, which is in accord with previous reports (Rommers et al., 2017; Wang et al., 2017; Terporten et al., 2019; see for a review, Prystauka & Lewis, 2019). In addition, and for the first time, the same pattern was evidenced in spoken comprehension, although the effect was less robust in this case (see below).

Modality-independent alpha power decreases during language prediction

Alpha power decreases were similar in both tasks in terms of spectral, temporal and spatial distribution. Specifically, both effects were most prominent in the alpha band, transient –starting and resolving during the delay–, and partially overlapped at frontal and central sites. Power decrease in alpha oscillatory activity during language prediction may be in agreement with the idea that desynchronization in the alpha and beta bands is inversely related to the amount of information to be encoded or retrieved from long-term memory (Hanslmayr et al., 2012). The main tenet of the 'information-via-desynchronization' hypothesis (Hanslmayr et al., 2012) is that an optimal brain state for engaging in cognitive processing might be during neural desynchronization (power decrease in alpha and beta oscillatory activity) (Griffiths et al., 2019). Considering this framework, language prediction might represent an information-rich scenario that requires larger cognitive engagement due to the faster retrieval of lexical-semantic features (see for a similar proposal in relation to the beta band, Molinaro et al., 2016). The sentence contexts in the HC condition allow the pre-activation of certain lexical or semantic features, which may be readily used to pre-update the ongoing high-level sentence meaning representation (Graesser, Singer & Trabasso, 1994; Singer, Graesser & Trabasso, 1994; Lau, Holcomb & Kuperberg, 2013), further incrementing the amount and the complexity of information that is being simultaneously processed. In contrast, in LC contexts, there is more uncertainty both at the level of the construction of the high-level sentence meaning representation and of the pre-activation of lexical-semantic representations, which therefore does not facilitate processing, as in the case of HC contexts. Following this interpretation, a larger amount of information to-be-retrieved and integrated might predict larger alpha power decreases, as it is observed in HC sentences in the present study.

In line with this interpretation, previous studies have shown that alpha-beta power is decreased in tasks that require greater semantic elaboration (Hanslmayr et al., 2009) and reduced upper alpha (10-12 Hz) over frontal sites has been associated with semantic processing demands (Klimesch et al., 1999; 1996). More specifically, some of these decreases in alpha-beta power have been observed in the left inferior prefrontal cortex, again pointing to its relevance for semantic processing (Hanslmayr et al., 2011; Meeuwissen, Takashima, Fernández & Jensen, 2011; Kielar, Panamsky, Links & Meltzer, 2015, but see Piai, Rommers, & Knight, 2018; Roos & Piai, 2020). Finally, a recent study found that alpha-beta modulations track the fidelity of stimulus-specific information represented in the cortex (Griffiths et al., 2019). Overall, the convergence across these studies and the present results reinforces our interpretation that language prediction might be accompanied by (modality-independent) alpha power decreases, reflecting larger processing demands due to the pre-activation and pre-updating of greater and richer information.

The current findings are also consistent with a more general view of decreases in alpha power as reflecting active engagement of relevant regions (Jensen & Mazaheri, 2010; Klimesch, 2012). In fact, as pointed by Griffiths and colleagues (2019), this account is also consistent with the previous interpretation, such that alpha and beta oscillations may boost or enable information processing through functional disinhibition of task-relevant networks. In line with this, a recent MEG study in written comprehension associated the pre-word alpha power decrease in HC (relative to LC) with increased activation of a brain network involved in high-level language processing (Wang et al., 2017). Furthermore, conjoint alpha-beta (8-25 Hz) power decreases are consistently found at left inferior parietal and temporal (and frontal) areas before naming a picture in HC contexts (e.g., Piai, Roelofs, Rommers & Maris, 2015; Piai, Rommers & Knight, 2018; Piai, Klaus & Rossetto, 2019; Roos & Piai, 2020), which is taken as a fingerprint of lexical and semantic retrieval during speech production (see for a review Piai & Zheng, 2019). The different scalp distribution and the absence of beta involvement in the current study may indicate that the effects observed here do not reflect the pre-activation (i.e. retrieval) of representations, but rather the operation of language-related or general-domain mechanisms that enable and support information processing during language prediction, as explained before.

Modality-dependent alpha power decrease in reading but not in spoken comprehension

On the other hand, the results also pointed to the involvement of modality-dependent operations. In the reading comprehension task, the alpha decrease extended to channels covering occipital sites (i.e. sensory-related regions), especially at a later stage in the pre-word interval (about –400 to –200 ms time-locked to the critical word). Pre-stimulus alpha power decreases in posterior sites are reliably found in anticipation of an imperative visual stimulus (Bidet-Caulet et al., 2012; Worden, Foxe, Wang, & Simpson, 2000; Thut, Nietzel, Brandt & Pascual-Leone, 2006) and often correlate with improved visual perception (e.g., van Dijk, Schoffelen, Oostenveld & Jensen, 2008; Jensen and Mazaheri, 2010; Jensen et al., 2012; Mathewson et al., 2014). Thus, it is possible that the posterior alpha power decrease observed here also reflected enhanced visual processing in HC (relative to LC). Participants may have treated the final word as a feedback stimulus – assuming they had generated a strong prediction in HC (perhaps at the level of the lexical form), they may have prepared to quickly test their prediction against the actual input and resume sentence integration as fast as possible.

In contrast, the alpha decrease in the spoken comprehension task did not suggest the involvement of modality-dependent areas. This could be nonetheless explained by the effect being overall less robust in this modality. Indeed, it was significant only after restricting the statistical test to a set of frontal electrodes that exhibited the alpha power decrease most prominently. There are several reasons that may account for a weaker effect in the auditory modality. First, the small size of the auditory cortices may lead to weaker and more blurred traces of alpha oscillations on the scalp level, which may explain why evidence for sensory alpha in the auditory domain is scarcer in general (for discussion, see Weisz, Hartmann, Müller, Lorenz & Obleser, 2011). Second, spoken sentences have variable durations, unlike sentences presented in written format, which all had the same duration. As a result, there may be a higher inter-trial variability in the timing of underlying neural processes, which, in turn, would lead to smaller effects in the averaged signal. Lastly, the method chosen for statistical testing may have been too conservative in this specific situation. Although the cluster-based permutation approach is appropriate for multidimensional data, it may be too conservative to detect small but meaningful effects (Groppe, Urbach & Kutas, 2011; Huang & Zhang, 2017). Despite these arguments, the finding in the spoken comprehension modality needs replication. In addition, given the current promising results, future studies could adopt a within-subjects design to improve statistical power and afford a more direct quantitative comparison between the two modalities.

Conclusion

In sum, the current study replicates previous reports of pre-word alpha power decreases in HC (relative to LC) consistent with prediction in written comprehension, and contributes in describing, although less robustly, the same pattern in spoken comprehension. Similarities in terms of spectral, temporal and spatial aspects of the observed alpha power decreases in both modalities of comprehension suggest the engagement of at least partially shared mechanisms. In particular, the findings are congruent with accounts of pre-stimulus alpha-beta decreases as modality-independent indices of information processing, in this case, pointing to greater and/or richer information processing during language prediction. The results are similarly consistent with greater engagement of areas involved in processing high-level aspects of predicted stimuli. Concurrently, the scalp distribution of the alpha power decrease suggests the involvement of modality-dependent operations to improve sensory processing in the reading, but not in the spoken comprehension task. Altogether, the current findings are in line with recent theorizing that emphasizes a role of predictive processing in language comprehension – when contexts afford the prediction of upcoming words, a series of mechanisms are engaged in advance to fine-tune the system for future bottom-up processing.

Study 4

Neural signatures of predictive language processing in Parkinson's disease with and without mild cognitive impairment

This study corresponds to:

León-Cabrera, P., Pagonabarraga, J., Morís, J., Martínez-Horta, S., Riba, J., Marín-Lahoz, J., Horta-Barba, A-, Bejr-Kasem, H., Kulisevsky, J., Rodríguez-Fornells, A. (2020) Neural signatures of predictive language processing in Parkinson's disease with and without mild cognitive impairment. *bioRxiv* 11.23.392647. DOI: https://doi.org/10.1101/2020.11.23.392647

8 Study 4

Neural signatures of predictive language processing in Parkinson's disease with and without mild cognitive impairment

8.1 Introduction

Parkinson's disease (PD) is a chronic, neurodegenerative disorder that, in addition to motor defects, involves difficulties in a variety of cognitive domains (Kudlicka et al., 2011; Muslimovic, Post & Speelman, 2005). Patients with PD may exhibit significant problems with language comprehension and language production in everyday life. These difficulties have been partly explained by studies that explored in-depth linguistic function in sentence comprehension (for a review, see Pell & Monetta, 2008) and have been mostly attributed to syntactic alterations (Lieberman, 1992; Friederici et al., 2002), although there is also evidence pointing to slower or delayed lexical and semantic activation (Arnott et al., 2001; Angwin et al., 2005; Angwin et al., 2017). Moreover, cognitive resource limitations in functions that enable or support language processing, such as executive functions or working memory, may secondarily impact comprehension (Grossman et al., 2002). Yet, there is still paucity of data in the literature on language processing and associated cognitive disturbances in non-demented PD patients.

An aspect of language comprehension that has not been evaluated in PD is predictive language processing, which plays an important role in achieving fast and efficient language processing (Kutas & Federmeier, 2011). Accordingly, readers and listeners probabilistically infer and pre-activate different aspects of upcoming words (e.g., van Berkum, Brown, Zwitserwoold, Kooijman & Hagoort, 2005; Wicha, Bates, Moreno & Kutas, 2003). Much evidence of language prediction has been derived from the N400 event-related potential (ERP) component, an index of semantic processing (Kutas & Hillyard, 1980). The N400, a negativity peaking ~400 ms after word onset, is reduced for words that are more predictable in a given context (N400 context effect) (Kutas & Hillyard, 1984), which is considered to reflect facilitated semantic processing owing to the prediction of the word or some of its semantic features (Federmeier & Kutas, 1999). Most relevantly, recent research has shown that processing differences arise even before the word is presented. In particular, slow negative potentials (SNP) consistent with semantic anticipation precede sentence-final words in predictive contexts (e.g., before "ball" in "The goalkeeper managed to catch the… ball") (León-Cabrera et al., 2017; 2019; Grisoni et al., 2017; for a review, see Pullvermüller & Grisoni, 2020).

Importantly, prior research suggests that language prediction is reduced in populations with limited cognitive resources (Federmeier et al., 2002; 2005; 2010; Payne & Federmeier, 2008; Wlotko and Federmeier, 2012; for reviews, see Payne & Silcox, 2019; Huettig, 2015). For instance, older adults seem to take less advantage of semantically rich contexts to facilitate subsequent processing (Federmeier and Kutas, 2005) and those with lower verbal fluency scores exhibit diminished or absent ERP effects associated with prediction, suggesting the adoption of a 'wait-and-see', incremental strategy instead (Federmeier et al., 2002). Crucially, up to 40% of individuals with PD will develop mild cognitive impairment (PD-MCI) within the first 5 years of disease (Aarsland et al., 2010; Litvan et al., 2011), which is a robust predictor of further conversion to dementia (PDD). Among the several cognitive phenotypes characterizing PD-MCI, changes in visuoperceptive skills, memory or language have been highlighted to be good predictors of the conversion from PD-MCI to PDD (Horta-Barba et al., 2020; Martínez-Horta & Kulisevsky, 2019; Lang et al., 2019). Based on this, we hypothesized that deficits in predictive language processing could contribute to sentence comprehension in PD, and that these early language deficits have a role in the development of PD-MCI.

To address language processing the current study investigated two ERP signatures of language prediction in PD with normal cognition (PD-NC) and PD-MCI. Following this, we presented sentence contexts that were predictive or not of a sentence-final word that was delayed by 1 s (see Table 2) (León-Cabrera et al., 2019), thus allowing to capture the abovementioned ERP signatures in the 'anticipatory' and 'processing' stages of prediction. In predictive contexts, half of the final words were semantically incongruent, thus representing a prediction mismatch. Lastly, we explored whether the targeted ERP signatures of prediction were reduced or diminished in PD patients with lower verbal fluency scores.

8.2 Methods

Participants. Participants were prospectively recruited from a sample of outpatients regularly attending the Movement Disorders Clinic at Hospital de la Santa Creu i Sant Pau (Barcelona, Spain). They were invited to participate in a longitudinal study involving exhaustive neuropsychological testing and two EEG recording sessions (one at baseline and a follow-up after 1 year). Healthy adults were also invited to participate in the study to serve as controls. The project was approved by the research ethics committee of the Hospital de la Santa Creu i Sant Pau. Before inclusion, written informed consent was obtained from all the participants.

Clinical and cognitive testing. PD patients had been diagnosed by a neurologist with expertise in movement disorders. All patients accomplished steps 1 and 2 of the UKPDSBB criteria, and three or more of the four first supportive positive criteria of step 3 (Hughes, Daniel, Kilford, & Lees, 1992). Motor status and stage of illness were assessed by the MDS-UPDRS-III and Hoehn & Yahr scales (Goetz et al., 2008). Demographic variables including age, gender, years of formal education, disease onset, medication history, as well as the levodopa equivalent daily dosage (LDD) (Tomlinson et al., 2010) were collected for all patients (see Table 1). All participants were classified as having either normal cognition or mild cognitive impairment (MCI) based on their score in the Parkinson's Disease - Cognitive Rating Scale (PD-CRS) (Pagonabarraga et al. 2008). A cutoff score of ≤ 83 was used to classify patients as with or without PD-MCI (Fernández de Bobadilla et al. 2013). Participants with dementia were excluded according to the MDS diagnostic criteria for PDD (Emre et al., 2007). In PD patients, cognition was examined during the 'on' state, and all participants were on stable doses of dopaminergic drugs during the 4 weeks before inclusion. In addition, their semantic and phonological verbal fluency was evaluated. For each participant, total verbal fluency scores were obtained by averaging the direct scores in semantic and phonological fluency subtests.

Final samples. A total of 135 participants completed the EEG recording session. Of those participants, 102 were PD patients and 33 were healthy controls. The data of some participants were excluded because the files were corrupt (n = 3) or their recordings were excessively noisy (n = 4) as determined by visual inspection of the raw data, leaving 128 participants available for analyses. From the remaining participants, 97 were PD patients and 31 were controls. Controls were excluded if they had mild cognitive impairment (n = 3) or their cognitive status was not available (n = 2). Lastly, participants with less than 20 available trials in any condition after EEG preprocessing (n = 9) were also excluded from the analyses (see EEG preprocessing section for more information about the artefact detection and rejection procedure).

The final pool of data consisted of 88 PD patients and 26 controls, from which matched samples were subsequently obtained to perform the following comparisons according to the goals of the present study: 1) PD patients with normal cognition (PD-NC) (N = 58) with an age-, gender- and education-matched Control Group (CG) (N = 24), and 2) a subsample of Parkinson's disease patients with normal cognition (sPD-NC) (N = 19) with an age-, gender-, education- and years of disease's evolution-matched PD patients with mild cognitive impairment (PD-MCI) (N = 20). To avoid the comparison of patients classified as sPD-NC or sPD-MCI but having similar PD-CRS total score, we selected participants with PD-CRS scores of < 75 for the sPD-MCI group and of > 85 for the sPD-NC.

The demographic and clinical features of the final samples are reported in **Table 1**. CG and PD did not differ in age (t(1,80) = –0.90, p = .37, 95% CI [–5.19, 1.95]), education (t(1,80) = 1.20, p = .23, 95% CI [–0.84, 3.43]), gender (X2(1, 82) = .853, p =.35), or verbal fluency scores (t(1,80) = 1.73, p = .08, 95% CI [–0.23, 3.53]). They differed in PD-CRS total scores (t(1,80) = 12.96, p <.001, 95% CI [26.2, 35.91]), but both groups were above the established score cutoff of 83 for PD-MCI diagnosis. Likewise, sPD-NC and sPD-MCI did not differ in age (t(1,37) = –1.23, p = 0.22, 95% CI [–3.89, .94]), education (t(1,37) = –1.30, p = 0.22, 95% CI [–1.30, 5.34]), gender (X2(1, 39) = .033, p =.855), disease's duration (t(1,37) = –0.53, p = .59, 95% CI [–2.84, 1.65]) or levodopa equivalent daily dose (t(1,37) = –1.74, p = .09, 95% CI [–293.94, 22.64]), but differed in PD-CRS total score (t(1,37) = 12.96, p <.001, 95% CI [26.2, 35.91]) and verbal fluency scores (t(1,37) = 6.59, p <.001, 95% CI [4.92, 9.29]). For all comparisons, equal variances were assumed based on null results in Levene's test.

Table 1. Demographics and clinical features of the patient and control samples.

	CG	**PD-NC**	P	**sPD-NC**	**PD-MCI**	P
N	24	58		19	20	
Age (years)	62.85 (6.82) [59.9, 65.7]	64.47 (7.61) [62.4, 66.4]	.37	70.75 (2.70) [69.4, 72]	72.22 (4.48) [70.1, 74.3]	.22
Education (years)	14.5 (4.04) [12.7, 16.2]	13.2 (4.57) [12, 14.4]	.23	12.42 (5.08) [9.9, 14.8]	10.4 (5.16) [7.9, 12.8]	.22
Women (%)	41.6	31	.35	42.1	45	.85
Disease duration (years)	n/a	5.26 (3.00) [4.4, 6]	-	5.68 (3.06) [4.2, 7.1]	6.28 (3.8) [4.4, 8]	.59
Levodopa dose (mg/d)	n/a	265.25 (249.58)	-	252.87 (293.74)	384.4 (215.04)	.12
LEDD (mg/d)	n/a	456.04 (230.68)	-	397.18 (236.08)	532.83 (209.70)	.09
PD-CRS total score [≤ 83 cutoff]	106.16 (8.74) [102.4, 109.85]	99.72 (10.86) [96.8, 102.5]	.01	98.2 (9.17) [93.7, 102.6]	67.15 (5.39) [64.6, 69.6]	<.001
Total verbal fluency (DS)	19.66 (3.12)	18.01 (4.18)	.08	17.63 (4.32)	10.52 (2.08)	<.001

The values correspond to means, standard deviations (in round brackets) and 95% mean confidence intervals (CI) (in square brackets). For simplicity, CI are shown only for the relevant variables in the group comparisons.

P-values were determined with independent samples t-tests between groups for continuous and ordinal variables and with Pearson Chi-Square for nominal variables (Gender).

n/a = not applicable; LEDD = Total levodopa equivalent daily dosage; PD-CRS = Parkinson's Disease Cognitive Rating Scale; DS = Direct score.

Materials. We used the materials from León-Cabrera et al. (2019) consisting of high (HC) or low (LC) constraining sentence contexts, establishing a stronger or weaker expectation for an upcoming word (see Table 2 for sentence examples). A total of 312 sentences were included, of which 208 had HC sentence contexts (66,6 %) and 104 had LC sentence contexts (33,3 %). The sentences were originally created and categorized as either HC (mean cloze probability = 76%, SD ± 17.7%) or LC (mean cloze probability = 6.1%, SD ± 10.3%) by Mestres-Missé, Rodríguez-Fornells and Münte (2007). Within the HC condition, half of the sentences ended with a congruent word and half with a semantically incongruent word. For the sake of task brevity to avoid fatigue, we decided to include this contrast (congruent vs. incongruent contrast) only in the HC condition. Sentences were further divided in two lists of 156 sentences, each list containing half of the sentences of each condition, and the use of one or another list was pseudorandomized across participants. In sum, each participant read the following sentences: 52 HC with congruent endings (HC congruent), 52 HC with incongruent endings (HC incongruent), and 52 LC with congruent endings (LC congruent) (Table 2). All final words were nouns. Congruent ones were the best completions – the word with the highest cloze-probability (i.e. percentage of individuals that supply that word as a continuation for that sentence) (Taylor, 1953) – and incongruent words were selected from the ESPAL database (Duchon et al., 2013) so that they matched the congruent endings in the variables of mean word length, mean number of syllables, word frequency, familiarity, imaginability and concreteness.

Table 2. Example sentences of each condition (original sentences in Spanish in brackets).

Condition	**Sentence context**	**Final word**
HC congruent	*The goalkeeper managed to catch the* (El portero fue capaz de atrapar la)	*ball* (pelota)
HC incongruent	*The goalkeeper managed to catch the* (El portero fue capaz de atrapar la)	*shore* (orilla)
LC congruent	*As a present she gave her son a* (Le ha regalado a su hijo una)	*ball* (pelota)

Procedure. Participants were tested individually. After external electrodes and the EEG cap were placed, electrode impedances were checked. Participants were instructed to keep fixation on the center of the screen and to blink only during the blinking interval. After completing a short training block (4 sentences), the proper experiment started. They were instructed to read attentively and silently the words that would appear at the center of the screen.

The experiment consisted of 26 blocks of 6 trials each (2 HC congruent, 2 HC incongruent, 2 LC congruent). The order of the sentences within each block was randomized. Words were white on black background (font type: Arial; font size: 24 points). On each trial, a fixation point (a cross) appeared at the center of the screen (800 ms). Then, a sentence was presented word-by-word (300 ms per word, followed by a 200 ms blank screen). Between the penultimate and the final word, the blank screen remained for a period of 1 s (pre-word interval). After the final word (300 ms duration), a blank screen was presented (1000 ms) before the blinking signal (1 s; depiction of an eye at the center of the screen). The next trial began after the offset of the blinking signal (see Figure 1 for a simplified display of the trial structure).

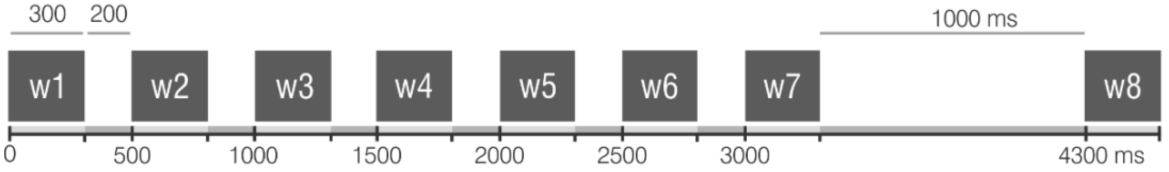

Figure 1. Depiction of the structure of a trial. Sentences were presented one word a time. Each word of the sentence context (w1 to w7) was displayed for 300 ms followed by a 200 ms inter-word interval, except for the last word (w7), which was followed by a 1000 ms interval (pre-word interval) leading up to the presentation of the final word (w8).

To ensure that participants were reading attentively, after every 2 blocks they were asked to judge whether the previous sentence was congruent or not ("did the previous sentence make sense?"). The question remained on the screen until they answered (yes/no) using the keyboard. The next block started immediately after the test except for even-numbered blocks, which were followed by a resting period that could be resumed anytime by the participant.

EEG recordings

EEG acquisition. Continuous electroencephalogram (EEG; sampling rate = 250 Hz) was recorded from 19 electrodes (FP1/2, F3/4, F7/8, Fz, C3/4, Cz, P3/4, Pz, T3/4, T5/6, O1/2) mounted on a cap following the international 10–20 system positions. The EEG signal was amplified on-line (band-pass filter = 0.016 – 1000 Hz) with Brain Vision. To monitor ocular artefacts, electrooculogram (EOG) electrodes were recorded with electrodes at dipolar vertical (supraorbital and suborbital ridge of the right eye) and horizontal (external ocular canthus of the left and right eyes) placements. All electrode impedances were kept below 5 kΩ. An on-line notch filter (50 Hz) was applied to attenuate high-frequency electrical noise and data were referenced to the mean activity of the left and right mastoid electrodes.

EEG preprocessing. EEG data were preprocessed using ERPlab v7.0.0 (López-Calderón & Luck, 2014) of the EEGlab toolbox v14.0.0 (Delorme & Makeig, 2004), FieldTrip version 20181231 (Oostenweld, Fries, Maris, & Schoffelen, 2011) and custom scripts programmed in MATLAB R2017b. For each participant and condition, the continuous data were segmented into the following epochs of interest: 1) from −100 to 1300 ms from the onset of the penultimate word (pre-word interval), and 2) from −100 to 1000 ms from the target word onset (word interval). Artefact detection and rejection were applied on the epoched data. To facilitate artefact rejection, we computed the vertical (vEOG) and horizontal (hEOG) bipolar EOG channels by subtracting the inferior from the superior vertical EOG channels and the right from the left horizontal EOG channels, respectively. For the pre-word interval, all epochs with activity ± 85 μV in the ocular channels or ± 200 μV in any other channel were automatically removed and the remaining epochs were visually inspected and excluded if they contained blinks, muscle activity or large drifts. For the word interval, the concatenated epochs were first subjected to independent component analysis (ICA) decomposition to correct blinks because preliminary visual inspection revealed that many participants had difficulties refraining from blinking until the blinking signal, leading to a large loss of trials in this interval. ICA was run using the "runica" algorithm as implemented in EEGlab toolbox v14.0.0, excluding the ocular channels from the decomposition. The ocular components were detected and rejected based on visual inspection. After artifact-correction, epochs with activity ± 200 μV in any but ocular channels were automatically rejected, and those that remained were visually screened for any previously unnoticed artefacts (e.g. muscle activity or large drifts).

For the CG with PD-NC comparison, an average of 89.8 % trials per condition were available for analysis (HC = 42.8 trials; LC = 41.9 trials; HCC = 49.6 trials; HCI = 49.5 trials; LCC = 49.7 trials; overall average = 46 trials per condition), with no significant differences in the percentage of available trials between groups (all p-values > 0.3). For the sPD-NC with sPD-MCI comparison, an average of 87.4% trials per condition were included (HC = 39.5 trials; LC = 39.2 trials; HCC = 49.3 trials; HCI = 49.4 trials; LCC = 49.5 trials; overall average = 45 trials per condition), with no significant differences in the percentage of remaining trials between groups (all p-values > 0.1).

Statistical analyses

Event-related potential analyses. Before performing statistical analyses, the data was baseline-corrected to the 100 ms pre-stimulus period and data were low-pass filtered (as implemented in ERPlab v7.0.0). For the pre-word interval, a 5 Hz low-pass Butterworth filter (12 dB/oct roll-off) was applied to the data to focus exclusively on slow activity (Brunia et al., 2011). For the word interval, in which the N400 component was targeted, a 30 Hz low-pass Butterworth filter (12 dB/oct roll-off) was used. For the figures presented, the data were filtered using a 12 Hz low-pass filter (roll-off of 12 dB/oct).

Cluster-based permutation analyses. Non-parametrical cluster-based permutation tests were used to assess the differences (Maris and Oostenveld, 2007) (as implemented in Fieldtrip version 20181231 under MATLAB R2017b), a method that exerts proper control for the increased probability of false positives in the context of many comparisons. Broadly, it identifies adjacent time points and channels with similar differences between conditions. The test worked as follows in the present study. For within-subject comparisons, every sample (channel x time) was compared between two conditions (e.g., HC vs. LC) through a dependent t-test. For between-subject comparisons, an independent t-test compared the effect (e.g., HC minus LC) between groups at each sample (channel x time). Next, adjacent samples were clustered based on a t-value threshold (pre-determined from a one-tailed t-distribution with an alpha level of .05 and N-1 degrees of freedom) and the cluster with the largest sum of t-values was selected (cluster-level t-value). To determine whether effects were significant, the Monte Carlo method was used to construct a null distribution from the cluster-level t-values of random partitions obtained by randomly swapping the samples (between conditions and within participants for within-subject comparisons; and between groups and within participants for between-subject comparisons) (5000 randomizations). Only observed clusters with cluster-level t-values within the 2.5th percentiles (alpha level of .05) of the null distribution were considered significant.

Choosing one-tailed testing is an optimal methodological decision when there are a priori hypotheses about the direction of the differences (Lakens, 2016). It allows to improve statistical sensitivity, which is particularly important when working with clinical samples. In this case, we expected more negative amplitudes for HC than LC in the pre-word interval (León-Cabrera et al., 2017; 2019). As for the interval after the presentation of the word, in which N400 context effects were targeted, we expected more negative amplitudes (larger N400 response) for unexpected that for expected words, that is, for HCI than for HCC, as well as for LCC relative to HCC (e.g., Kutas & Federmeier, 2011).

Correlations between ERP effects and verbal fluency. After examining ERP effects, we conducted non-parametrical Spearman-Brown correlations to evaluate a potential linear relationship between the ERP effects and verbal fluency performance. All PD patients were included in this analysis (N = 78), that is, PD-NC (N = 58) and PD-MCI (N = 20). For each participant, the total verbal fluency score was obtained by averaging the direct scores in semantic and phonological fluency subtests (see Supplementary Materials for correlations with each subtest individually). ERP measures were labeled and quantified as follows: 1) SNP (LC minus HC mean amplitude difference in the 600 ms before pre-word onset at a left anterior cluster including electrodes FP1, F3, Fz, F7); 2) N400 congruency effect (HCC minus HCI mean amplitude difference from 300 to 500 ms post-word onset); 3) N400 constraint effect (LCC minus HCC mean amplitude difference from 300 to 500 ms post-word onset); and, based on the ERP results, 4) prolonged N400 congruency effect (HCC minus HCI mean amplitude difference from 600 to 800 ms post-word onset). All effects tied to the N400 were tested on a centrally distributed cluster (averaged electrodes: Cz, C3, C4, Pz, P3, P4). All values were normalized before performing correlations. All p-values were corrected with the Holm-Bonferroni correction (Holm, 1979).

8.3 Results

Event-related potentials

CG vs. PD-NC

Pre-word interval. Figures 2A and 2B show the grand-average ERPs at the F7 electrode during the pre-word interval in the CG group and the PD-NC group, respectively. The ERPs were computed in the interval from the onset of the penultimate word to the onset of the final word (1300 ms duration, the first 300 ms correspond to the penultimate word, and the last 1000 ms to the delay interval), with a 100 ms pre-stimulus baseline.

Both groups exhibited similar patterns, namely, an N100 associated with the processing of the penultimate word, followed by a positivity, and the later development of a negativity that became progressively larger as the presentation of the final word approached. As expected, based on previous findings in healthy adult population (León-Cabrera et al., 2019), the negativity was more prominent in the HC condition (relative to the LC condition), that is, when the final word could be strongly predicted from the prior context. This difference was confirmed by a significant cluster in the CG group ($p = 0.01$) (Fig. 2A) as well as in the PD-NC group ($p = 0.01$) (Fig. 2B). The cluster showed a similar temporal profile in both groups. It started approximately about 800 ms before the final word and went on until the word appeared. In regard to the spatial distribution, the cluster encompassed mainly frontal electrodes, with a slight left lateralization. There were no significant differences between groups.

Final word interval. We then turned to the interval after word presentation. As can be observed in Figures 2C and 2D, for both groups, unexpected words (HCI and LCC) elicited a larger negativity peaking about 400 to 600 ms (relative to expected words; HCC) that is consistent with the canonical features of the N400. These differences were reflected in the following significant clusters.

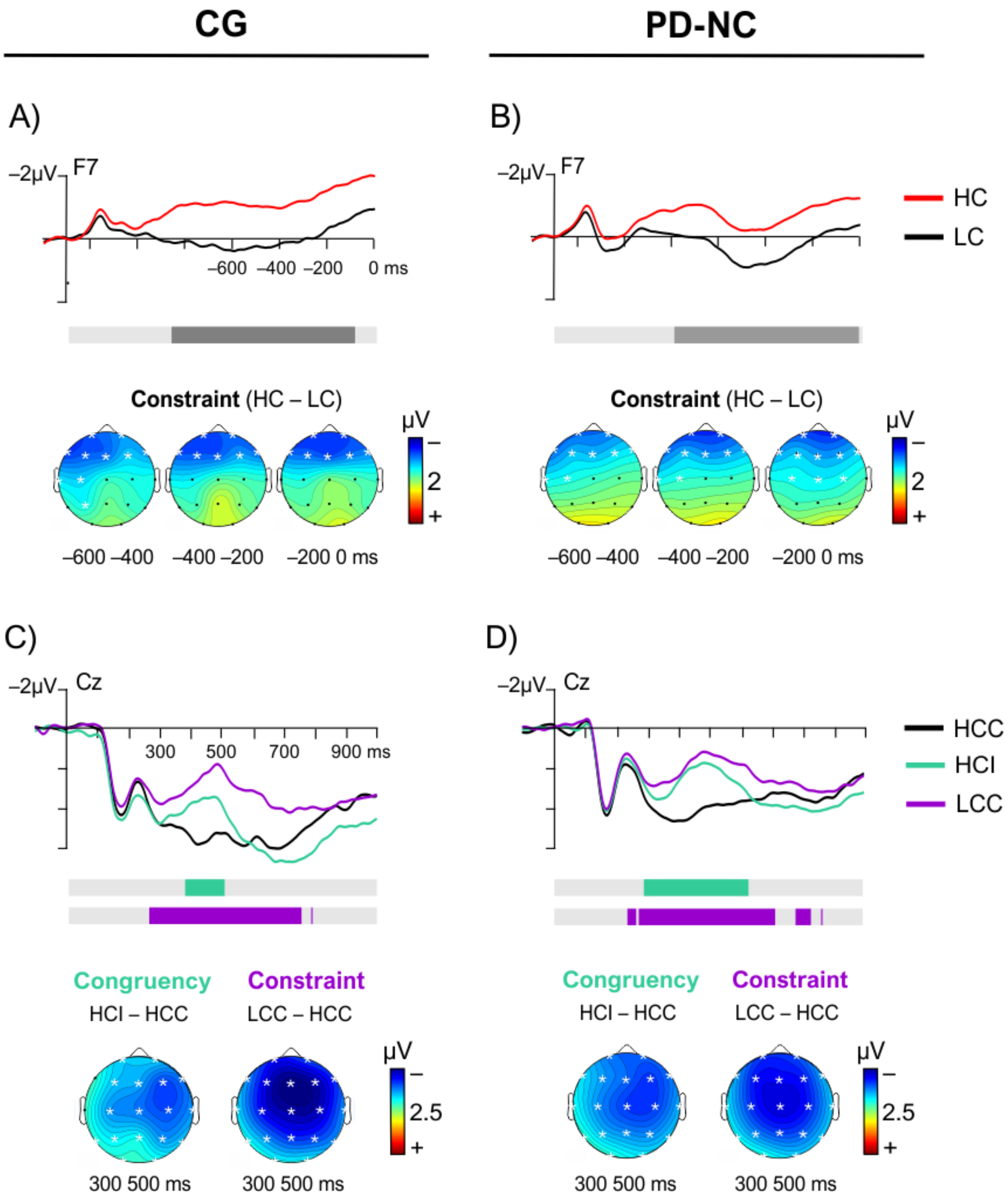

Figure 2. ERP results for CG and PD groups. A and B) Grand-averaged ERPs at the F7 electrode during the pre-word interval for both conditions for **(A)** CG and **(B)** PD. ERPs are time-locked to the presentation of the penultimate word (100 pre-stimulus baseline). The horizontal bar highlights (grey) the time interval in which significant differences were obtained. Below the ERPs, topographical maps indicating the electrodes (white asterisks) that showed a significant effect of constraint (HC minus LC) in the 600 ms prior to the critical word (in steps of 200 ms). **C and D)** Grand-averaged ERPs at Cz electrode after the presentation of the sentence-final word (100 pre-stimulus baseline) for **(C)** CG and **(D)** PD. The horizontal bars highlight the time interval of the N400 effect of Congruency (green) and Constraint (purple). Topographical maps with electrodes (white asterisks) that showed effects in the time interval in which N400 effects were expected to be maximal (300 to 500 ms post-word onset).

The effect of constraint (HCC versus LCC) was captured by significant clusters in the CG group (p = .001) (Fig. 2C) and in the PD-NC group (p <.001) (Fig. 2D), whereby words presented in less predictive contexts (LCC) showed more negative amplitudes than words in strongly predictive contexts (HCC). There were no differences between groups. The cluster started about 250-300 ms after word onset and resolved at 700 ms, and it exhibited a widespread distribution over the scalp. In fact, the cluster was significant at all sites. On the other hand, the effect of congruency (HCC versus HCI) was reflected in a shorter-lived cluster in the CG group (p = .007) (Fig. 2C) and PD-NC group (p<.001) (Fig. 2D), in this case, confirming more negative amplitudes for incongruent than congruent words in strongly predictive contexts. Again, there were no significant differences between groups. The effect was widespread with a seemingly slight rightward focus in magnitude.

sPD-NC vs. PD-MCI

Pre-word interval. The mean voltage pattern elicited during the pre-word interval (Fig. 3A and Fig. 3B) exhibited the normative N100 component, followed by a transient positivity and a subsequent negativity in the last part of the interval that evidenced the hypothesized tendency towards more negative amplitudes for HC than LC. However, in this case, there were no statistically significant differences between conditions, nor any differences between the two groups.

Final word interval. After the presentation of the final word (Fig. 3C and Fig. 3D) there was an observable larger negativity for unexpected (HCI and LCC) relative to expected words (HCC) in both groups, in line with classical N400 context effects. These observable differences were confirmed by significant clusters as specified next.

There were significant effects of constraint (HCC versus LCC) both for the group of patients without MCI (p <.001) (Fig. 3C) and the group with MCI (p <.001) (Fig. 3D). That is, as expected, words in less predictive contexts elicited a larger N400 than words in strongly predictive contexts, a difference that encompassed the interval between about 300 to 700 ms post word onset. The scalp distribution of the effect covered a wide set of electrodes over frontal and central regions of the scalp, with a slight frontward amplitude maximum. No differences between groups were found.

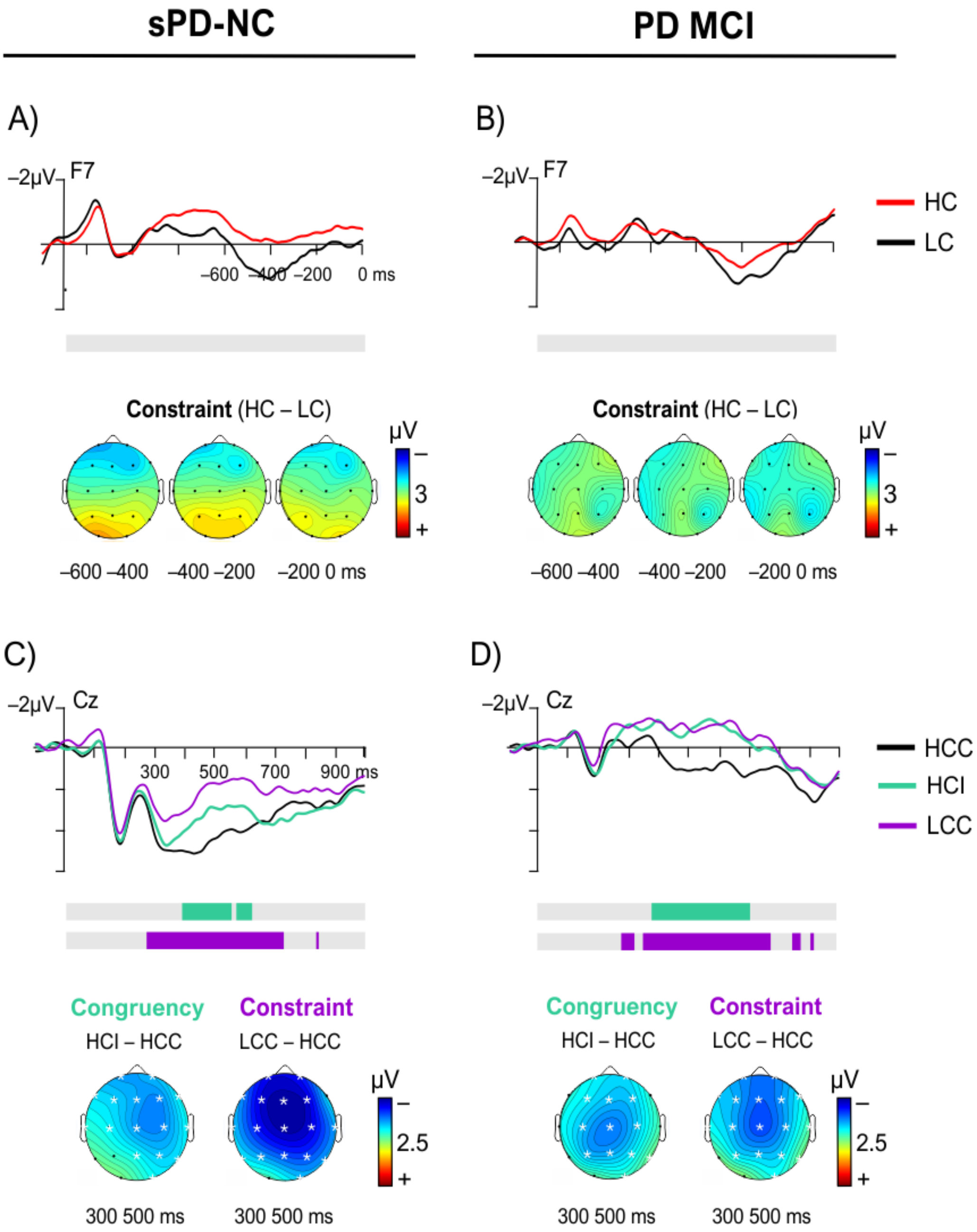

Figure 3. ERP results for sPD-NC and PD-MCI groups. A and B) Grand-averaged ERPs at the F7 electrode during the pre-word interval for (A) sPD-NC and (B) PD-MCI. ERPs are time-locked to the presentation of the penultimate word (100 pre-stimulus baseline). The horizontal bar highlights (grey) the time interval in which significant differences were obtained. Below the ERPs, topographical maps indicating the electrodes (white asterisks) that showed a significant effect of constraint (HC minus LC) in the 600 ms prior to the critical word (in steps of 200 ms). **C and D)** Grand-averaged ERPs at Cz electrode after the presentation of the sentence-final word (100 pre-stimulus baseline) for (C) sPD-NC and (D) PD-MCI. The horizontal bars highlight the time interval of the N400 effect of Congruency (green) and Constraint (purple). Topographical maps with electrodes (white asterisks) that showed effects in the time interval in which N400 effects were expected to be maximal (300 to 500 ms post-word onset).

As for the effect of congruency in predictive contexts (HCC versus HCI), again as expected, significant clusters reflected that incongruent words had more negative amplitudes than congruent words in patients without MCI (p = .005) (Fig. 3C), and also in with MCI (p = .001) (Fig. 3D). Relevantly, in this case, there were significant differences between the groups (see Figure 4). The difference waveforms (HCI minus HCC) of each group were contrasted in a between-group cluster-based permutation test including all electrodes and time-points (from 0 to 1000 ms). The test yielded a significant cluster (p = .008) that revealed a longer-lasting effect of congruency for the group of patients with MCI, which can be visualized in Figure 4A. More specifically, the cluster spanned approximately from 630 to 760 ms after word onset at centro-parietal sites (significant electrodes: C3, Cz, C4, P3, Pz, P4, T5, T4, T6, O2) (see the scalp distribution of t-values in Fig. 4B). This result suggests a prolonged N400 congruity effect in the PD-MCI group relative to the group of patients without MCI.

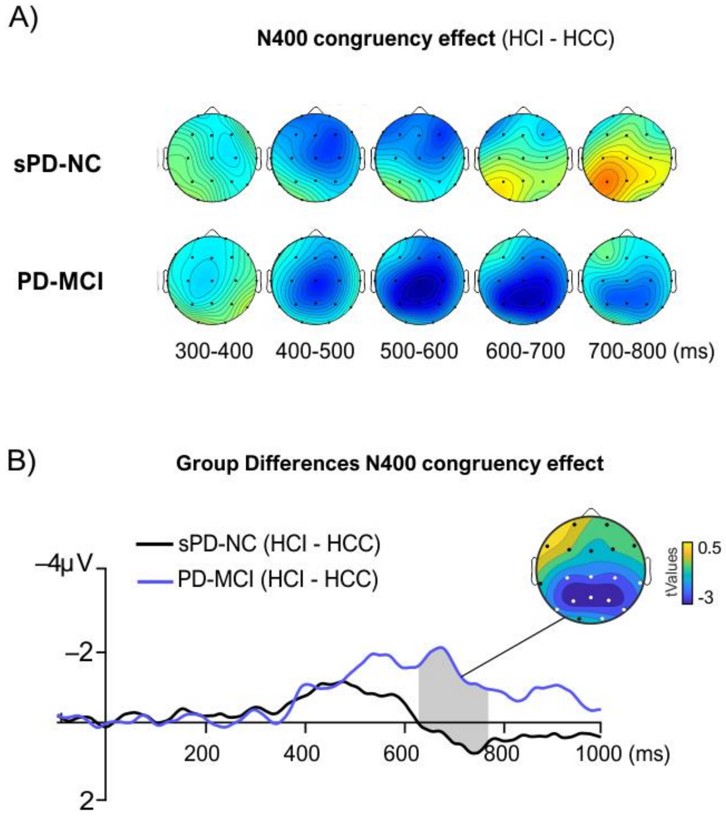

Figure 4. A) Topographical maps showing the temporal evolution and scalp distribution of the congruency contrast (HCI minus HCC) during final word processing, showing a prolonged N400 congruency effect in the group of individuals with PD and MCI. **B**) Grand-averaged ERPs of the difference waveforms (HCI and HCC) of both groups, averaged over the set of electrodes that were part of the significant cluster of differences between the groups (highlighted in white in the topographical map). The grey-colored area shows the timepoints included in the cluster. The topographical map shows the scalp distribution of the averaged t-values within the cluster (cluster t-value ± 2.02 for an alpha level of .05).

Correlations between ERP effects and verbal fluency scores

After establishing the ERP effects, we examined the association between ERP measures and verbal fluency scores in all PD patients (N = 78) (for further specification of this analysis, see Section 3.2) by means of non-parametric Spearman correlations between direct scores in verbal fluency tests, and four ERP measures, namely, 1) SNP (LC minus HC), 2) N400 constraint effect (LCC minus HCC), 3) N400 congruency effect (HCC minus HCI), and 4) prolonged N400 congruency effect (HCC minus HCI). Note that, in these analyses, the direction of the condition subtraction is reversed, such that positive values represent differences in the expected direction of the effect (e.g., in the case of the SNP, positive values indicate larger amplitudes for HC than LC). All correlations were corrected using Holm-Bonferroni correction (indicated as pHB in the text).

Table 3. Results of correlational analyses between ERP measures and verbal fluency scores.

Test	**ERP effect**	**r**	***p*-value (corrected)**
Verbal fluency	SNP	.326	**.010***
	N400 constraint	.157	.336
	N400 congruency	−0.073	.525
	Prolonged N400 congruency	–.358	**.005***

SNP (LC minus LC); N400 constraint (LCC minus HCC); N400 congruency (HCC minus HCI); Prolonged N400 congruency (HCC minus HCI).

Holm-Bonferroni correction was applied to adjust p-values. Statistically significant correlations are highlighted with asterisks.

The results of the correlational analyses are presented in Table 3. After correcting for multiple omparisons, two of the ERP measures showed a significant correlation with verbal fluency scores: the SNP and the prolonged N400 congruity effect. In particular, the SNP showed a significant positive correlation with verbal fluency (r(77) = .326, pHB = .010). Note that we excluded one participant who had outlying scores in the SNP (z-score = −3.91) although the correlation was unaffected by its inclusion (r(78) = .326, pHB = .01), which is normal when using nonparametric Spearman correlations that are very robust to the presence of outliers. This indicates that a larger SNP (more negative amplitudes prior to words in predictive than in unpredictive contexts) is associated with better verbal fluency performance in patients with PD. In turn, the prolongation of the N400 congruency effect (more negative amplitudes for incongruent than congruent words 600-800 ms post-word) correlates with worse scores in verbal fluency (r(78) = –.358, pHB = .005). No other correlations were statistically significant.

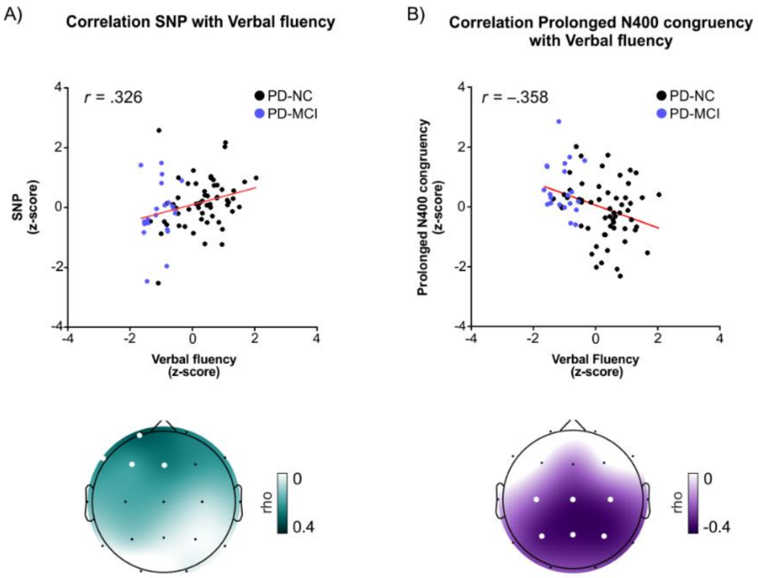

Figure 5. Scatterplots and scalp maps of significant correlations between ERP measures and verbal fluency scores in all PD patients (PD-NC and PD-MCI) (N = 78). A) Scatterplot of the positive correlation (r = .326, pHB = .010) between verbal fluency scores and the SNP (more negative amplitudes prior to words in HC than in LC contexts) and **B)** of the negative correlation (r = –.358, pHB = .005) between verbal fluency scores and the prolongation of the N400 congruency effect (more negative amplitudes for incongruent than congruent words 600-800 ms post-word). All values were normalized. Below the scatter diagrams, the r-values (Spearman rho) for all 19 electrodes are plotted on an idealized head. Darker shading indicates larger r-values in the direction of the correlation. Electrodes included in the ROI used to perform the reported correlations with verbal fluency (shown in the scatter diagrams) are highlighted (in white). R-values between electrodes were estimated by spherical spline interpolation.

8.4 Discussion

Motivated by previous evidence of altered predictive language processing in populations with low cognitive resources (e.g., Federmeier et al., 2002; 2005; 2010), the present study assessed language prediction in PD with normal cognition (PD-NC) and in PD with concomitant MCI (PD-MCI). To this end, ERP modulations were examined before and after encountering words that could or could not be predicted from context, thus capturing both the anticipatory and processing stage of prediction. Specifically, we focused on SNPs as signatures of semantic anticipation, and post-word N400 context effects as indices of predictive pre-activation. Relative to controls, PD-NC patients exhibited expected ERP signatures, pointing to preserved predictive language processing. In turn, PD-MCI patients showed absent SNPs and a significant prolongation of N400 congruency effects in predictive contexts, suggesting altered mechanisms tied to language prediction. Interestingly, correlational analyses revealed that worse verbal fluency performance was associated with the presence of the ERP pattern observed in the PD-MCI group.

Neural correlates of predictive sentence processing are preserved in PD-NC

Consistent with prediction, PD-NC patients exhibited the canonical reduction of N400 amplitude for expected (HCC) compared to unexpected words (LCC and HCI) (Kutas & Hillyard, 1984). This effect is taken to reflect facilitated lexical and semantic activation of expected words as a result of their pre-activation (Kutas & Federmeier, 1999; Lau, Phillips, & Poeppel, 2008). In line with this interpretation, in predictive contexts (HC), PD-NC individuals exhibited a frontally distributed negative SNP that developed progressively prior to the presentation of the final word (León-Cabrera et al., 2017; 2019; for similar results, see Grisoni et al., 2017; Li et al., 2017). Studies in healthy population suggest that such SNPs capture anticipation of semantic aspects of the upcoming word (for a recent review, see Pullvermüller & Grisoni, 2020). More specifically, it may reflect general-domain of language-specific mechanisms supporting retrieval or maintenance of pre-activated representations (Li et al., 2017; León-Cabrera et al., 2019). Altogether, PD-NC patients seem to make proper use of sentence contexts, affording normal semantic processing in PD, as has been previously observed (Friederici et al., 2002). In addition, the current study shows, for the first time, that they can also use sentence contexts to anticipate and pre-activate upcoming information.

Previous N400 studies have pointed to altered lexical and semantic activation in PD-NC (Angwin et al., 2017; Kutas et al., 2013; Angwin et al., 2004; Arnott et al., 2010, Copland et al., 2009; Angwin et al., 2007). However, these studies employed single-word contexts, instead of sentence contexts. Importantly, sentence contexts involve the construction of a high-level meaning representation (Graesser, Singer & Trabasso, 1994), which provides additional constraints to pin down relevant lexical and semantic representations, perhaps compensating for potential activation problems within semantic networks. Nonetheless, note that all patients in the current study were on dopaminergic medication. PD-NC patients on dopaminergic medication perform better in semantic priming tasks than PD-NC patients off medication (Angwin et al., 2004a; Angwin et al., 2009), suggesting that dopamine (DA) mediates semantic activation, arguably through calibration of the spread and focus of activation within semantic networks (Kischka et al., 1996). DA has also been linked with anticipatory processes in PD. Specifically, PD-NC patients off medication show reduced amplitudes of the stimulus-preceding negativity (SPN), suggesting that the DA system is implicated in the anticipation of motivationally salient and rewarding stimuli (Mattox, Valle-Inclán & Hackley, 2006). Future studies testing patients off medication would help untangle to what extent dopaminergic compensation contributed to the preservation of otherwise altered mechanisms.

PD-MCI impacts late semantic processing in predictive contexts

To further investigate the impact of cognitive impairment, we also examined neural responses in a smaller sample of PD-MCI patients. Compared to matched PD-NC patients (sPD-NC), PD-MCI patients exhibited expected N400 amplitude modulations within the normal onset latency (i.e., about 300-400 ms post-word onset). However, most remarkably, PD-MCI patients showed a significantly prolonged N400 congruency effect in predictive contexts, extending up to 700 ms after word onset. Critically, this was not observed in the N400 constraint contrast (LCC versus HCC), which suggests that the prolongation stemmed from the ERP response to incongruent words, rather than to congruent words (i.e. from HCI, instead of HCC). With that in mind, it is important to note that N400 context effects do not solely capture prediction effects, but also integration demands (van Berkum, Brown & Hagoort, 1999). Interestingly, a recent study showed that prediction effects dominate the earlier portion of the N400 context effect (starting as early as 200 post-word onset), whereas integration effects began later and continued until about 650 ms after word onset (Nieuwland et al., 2020; see also, Brothers et al., 2014; Lau et al., 2016). As such, prolonged N400 congruency effects may reflect abnormally sustained efforts to integrate the incongruent word with the previous context (Nieuwland et al., 2020). Curiously, a previous study found that patients with bilateral basal ganglia lesions exhibited similarly prolonged N400 effects in sentence processing (up to 700 ms post-word) (Kotz et al., 2003). In PD, the worsening of the condition from PD to PD-MCI may lead to altered late integrational semantic processing as well.

More recently, N400 effects have postulated to reflect the ‘prediction error’ (albeit with nuances in their definition, see Willems, Frank, Nijhof, Hagoort & van den Bosch, 2016; Rabovsky, Hansen, & McClelland, 2018; Kuperberg, Brothers, & Wlotko, 2020). In predictive processing frameworks (Clark, 2013), predictions are sent top-down from higher to lower level regions to ‘explain away’ the incoming sensory input and, then, the ‘prediction error’ – the portion of information that remains unexplained – is sent back and used as a learning signal to update internal models and improve future predictions. Thus, prolonged N400 congruency effects may indicate difficulties in this ‘learning’ phase, leading to less efficient internal model updating. In line with less efficient predictive language processing in PD-MCI, there were was no SNP in this group, pointing to absent semantic anticipation. In this case, however, the effect was absent in the sPD-NC group as well, and therefore cannot be attributed to MCI.

Worse verbal fluency performance is associated with altered signatures of predictive language processing in PD

Interestingly, correlational analyses revealed that lower verbal fluency scores were associated with reduced SNPs and prolonged N400 congruency effects. Similarly, previous evidence has shown that verbal fluency scores are a good predictor of the status of predictive language processing in healthy older population (Federmeier, McLennan, De Ochoa-Dewald & Kutas, 2002). Worse performance in verbal fluency is generally associated with executive dysfunction, such as difficulties in rule switching or inhibition. Executive dysfunction is common with the progression of fronto-striatal deterioration in PD and in PD-MCI (Kudlicka, Clare & Hindle, 2011; Monchi et al., 2004; Aarsland et al., 2010) and has been proposed to indirectly hinder language processing in PD (Grossman et al., 2002; 2003). However, verbal fluency performance depends not only on frontal lobe function, but also on lexical and semantic retrieval processes dependent on temporo-parietal structures (Unsworth, Spillers & Brewer, 2011). In fact, recent findings point to a greater weight of language-related processes than executive function in verbal fluency (Whiteside et al., 2016). Recent meta-analyses and lesion-based studies have also shown that verbal fluency is supported by standard language networks and underlying white-matter connectivity (Griffis et al., 2017; Costafreda et al., 2006; Baldo et al., 2006). Therefore, the ERP pattern associated with lower verbal fluency – suggestive of less efficient predictive processing mechanisms, as previously dicussed – may reflect not only executive dysfunction, but also disruption in semantic networks, in line with recent findings of difficulties in temporal-dependent functions in PD and PD-MCI (Horta-Barba et al., 2020; Martínez-Horta & Kulisevsky, 2019; Lang et al., 2019).

Concluding remarks and future directions

Overall, the results suggest preserved predictive language processing in PD-NC. In turn, in PD-MCI, further cognitive limitations hinder mechanisms associated with semantic prediction in normal circumstances. While these limitations may not prevent the pre-activation of relevant representations, they might negatively affect sentence processing due to poorer semantic anticipation and less efficient semantic processing when contexts are predictive. Furthermore, both executive dysfunction and damage within semantic networks may underlie these difficulties, as suggested by their association with low verbal fluency scores. Along with recent research (Horta-Barba et al., 2020; Martínez-Horta et al., 2019; Lang et al., 2019), these findings emphasize the value of examining cognitive changes in PD and PD-MCI in domains beyond executive functions, like the language domain. Interestingly, altered N400 effects are associated with a higher risk of transiting to dementia in individuals with amnesic MCI (Olichney et al., 2008). Therefore, future longitudinal studies in PD could evaluate changes in N400 congruency effects as potential markers of the transition to PD-MCI and PDD (for interested readers, see Supplementary Materials for a hierarchical regression analysis further attesting that the N400 congruency effect is a good predictor of global cognitive decline in PD). Finally, as previously highlighted, future studies testing patients off medication could assess the exact role of DA in predictive sentence comprehension. Yet, the value of evaluating PD patients on stable doses of dopaminergic medication must be emphasized, as it is representative of their habitual cognitive status and thus has the potential to uncover limitations that may impact the everyday functioning of individuals with PD.

8.5 Supplementary materials

Correlational analyses between ERP measures and verbal fluency (semantic and phonological separately)

To provide a more complete picture of the relation between the ERP measures and verbal fluency performance, we additionally performed a correlational analysis with semantic and phonological fluency separately. The parameters used to compute the ERP measures were the same as those used for the analyses in the main paper (see Section 3.2 in the Methods). For each group, the mean and standard deviation of the scores in semantic and phonological fluency tests are presented in **Table** 1.

Table 1. Demographics and clinical features of the patient and control samples

	CG	**PD-NC**	P	**sPD-NC**	**PD-MCI**	P
N	24	58		19	20	
Semantic fluency (DS)	22.08 (3.67)	20.03 (5.10)	.07	19.63 (6.29)	12.05 (3.57)	<.001
Phonological fluency (DS)	17.25 (3.99)	16 (4.9)	.27	15.63 (4.16)	9 (3.22)	<.001

The results of the correlational analysis are shown in **Table 2.** Two correlations survived after correcting for multiple comparisons (using Holm-Bonferroni correction, indicated as pHB), which were essentially the same as in the main paper, but only with phonological fluency in this case. Specifically, phonological fluency scores correlated positively with the SNP (r(78) = .393, pH = .003) and a negatively with prolonged N400 effects (r(78) =–371, pHB = .005) (see **Figure** 1). Note, however, that the correlations between these ERP measures (SNP and prolonged N400 congruency effects) and semantic fluency scores were significant or marginally-significant before correcting (p = 056 with the SNP; and p = .01 with the N400 congruency effect). Indeed, the correlation with the N400 congruency effect was marginally non-significant even after correction (p = .063). Likewise, the correlation between semantic fluency and the N400 constraint effect remained marginally non-significant after correction (p = .063).

Table 2. Results of correlational analyses between ERP measures and verbal fluency scores (semantic and phonological fluency).

Test	ERP effect	r	*p*-value	Test	ERP effect	r	*p*-value
Semantic fluency	SNP	.217	.056 (.225)	**Phonological fluency**	SNP	.393	**<.001* (.003*)**
	N400 constraint	.287	**.010*** (.063)		N400 constraint	–.020	.856 (1.713)
	N400 congruency	-.007	.949 (1.713)		N400 congruency	–.118	.300 (.902)
	Prolonged N400 congruency	-.284	**.011*** (.063)		Prolonged N400 congruency	–.368	**<.001* (.006*)**

SNP (LC minus LC); N400 constraint (LCC minus HCC); N400 congruency (HCC minus HCI); Prolonged N400 congruency (HCC minus HCI)

Corrected p-values (Holm-Bonferroni correction) are shown in brackets. Statistically significant correlations are highlighted with asterisks.

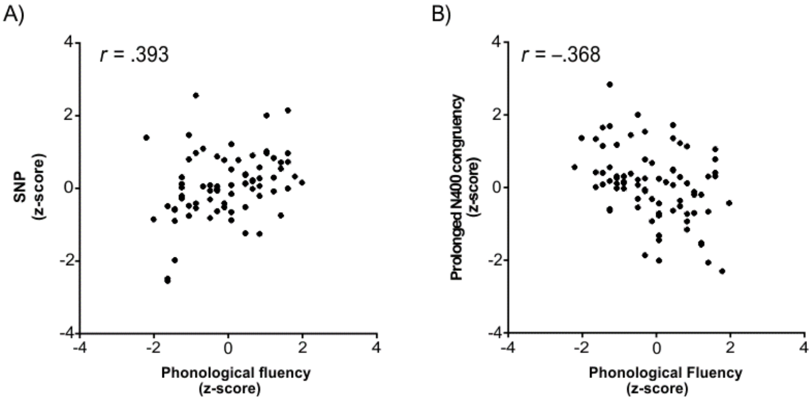

Figure 1. Scatterplots of significant correlations between ERP measures and phonological verbal fluency in all PD patients (with and without MCI). A) Scatterplot showing the significant positive correlation between phonological fluency scores and the SNP (more negative amplitudes prior to words in predictive than in unpredictive contexts). All values were normalized. B) Scatterplot of the significant negative correlation between phonological fluency scores and a prolongation of the N400 congruency effect (more negative amplitudes for incongruent than congruent words 600-800 ms post-word). All values were normalized.

Linear regression between PDCRS scores and prolonged N400 congruency effects

Based on the finding of a prolonged N400 congruency effect, we further investigated if this effect was a good predictor of global cognitive function, operationalized as the PD-CRS total score, in the whole sample of individuals with PD (N = 78), controlling for age and education.

For each participant, we quantified the prolonged N400 congruency effect (hereafter referred to as Late N400) as the mean amplitude difference between HCI and HCC in all electrodes that showed a significant difference in the exploratory analyses: C3, Cz, C4, P3, Pz, P4, T5, T4, T6, O2) in the 600 to 800 ms interval (time-locked to final-word onset). Note that the electrodes used to compute the prolonged N400 effect here are different from those used in the main text (C3, Cz, C4, P3, Pz, P4). This is because the current analysis was tailored for the prolonged N400 congruency effect exclusively, whereas, in the main text, we used the same electrodes for all N400 effects (i.e., congruency and constraint) for parsimony and comparability.

A two-stage hierarchical multiple regression model was performed to predict the PD-CRS using age and education at the first stage and adding the prolonged N400 congruency effect at the second stage (**Table 3**). This procedure allowed to evaluate the unique contribution of the ERP measure to explain global cognitive performance while controlling for age and education. The first model accounted for 27% of the variance (R^2 = .275; F(2,77) = 14.24, p < .001) and both age and education were significant predictors of the PDCRS total score (both p < .03). The addition of the Prolonged N400 measure measure as a predictor significantly improved the model (F_{change}(2,75) = 10.25, p = .002), explaining 36 % of the PD-CRS score (R^2 = .363; F(3,74) = 14.08, p < .001) (**see Fig. 2**). All predictors were significant in the second and final model (all p-values < .03). All assumptions for multiple linear regression were met. There were no problems of multicollinearity (all tolerance values were higher than .2 and the variance inflation factor was higher than 10) (Menard, 1995; Myers, 1990), or outliers or influential cases (all cases had Cook's distance > 1 and Mahalannobis values > 15) and residuals were equally and normally distributed (assessed through plot inspection).

Table 3. Coefficients for the two-stage hierarchical multiple regression with PDCRS as a dependent variable and age, education (stage 1) or age, education and prolonged N400 congruency effect (stage 2) as predictors.

	Unstandardized coefficients		**Standardized coefficients**		
Model 1	**B**	**SE**	**β**	**t**	**p**
(Constant)	135.33	18.49		7.31	< .001
Age	−.830	.241	−.370	−3.45	.001
Education	.897	.382	.252	2.35	.021
Model 2					
(Constant)	125.52	17.71		7.08	<.001
Age	−.652	.234	−.291	−2.78	.007
Education	.850	.361	.239	2.35	.021
Prolonged N400 congruency effect	2.2	.693	.309	3.20	.002

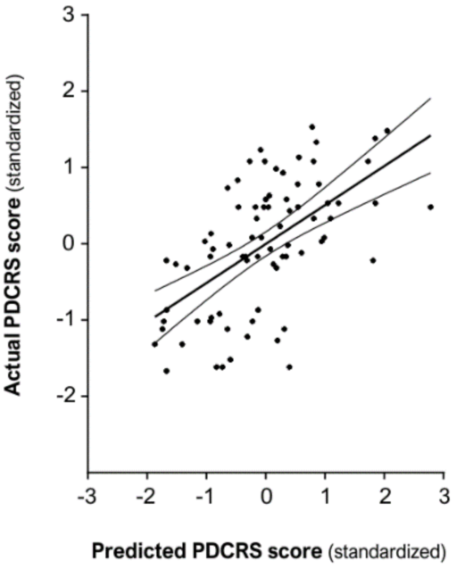

Figure 2. Scatterplot of observed PDCRS scores (Y axis) by predicted PDCRS scores (X axis) obtained in the linear regression model with age, education and the Prolonged N400 congruency effect (HCI minus HCC difference waveform in the 600-700 ms over a centro-parietal cluster), capturing the prolongation of the N400 congruity effect, as predictors. Note that smaller values of the Prolonged N400 congruency effect correspond to larger N400 effects (i.e. the subtraction between HCI and HCC yielding a negative value, indicating more negative amplitudes for incongruent than congruent words). Scores are standardized. A linear regression line was fitted to the data with 95% confidence intervals.

General Discussion

9 General Discussion

Until now, most studies of predictive language processing have focused on what occurs after linguistic predictions about certain words are confirmed or disconfirmed by the input. Here, we developed a novel approach to explore neural indices of mechanisms that support prediction in language *before* critical words are perceived. Although the findings of each study have been discussed in detail in each chapter, I will next provide an overview of the insights gathered from this work and attempt to relate them more generally with prior knowledge.

9.1 What occurs before finding out? Neural indices of semantic anticipation in sentence comprehension

9.1.1 Left anterior slow negative potentials

Based on an abundant corpus of knowledge demonstrating that pre-stimulus SNPs index anticipatory behaviour in other cognitive domains, we reasoned that differences in anticipation due to semantic predictions could be associated with ERP. To test this idea, we developed a novel paradigm to examine the anticipatory stage of language prediction. We presented sentence contexts that established different levels of expectancy about a final word. Crucially, in order to explore anticipatory processes, we inserted a short break (1 s) before the final word. We described left anterior SNPs (i.e., slow negative potentials) during this break with an amplitude that is associated with the level of contextual constraint, reaching maximal differences before the presentation of the final word. We also tested whether this effect was dependent or independent of comprehension modality, and in line with modality independence, we found the same pattern in spoken (**Study 1**) and in reading (**Study 2**) comprehension, at the same time demonstrating the robustness of the effect. In addition, to obtain a full picture of sentence processing in the same task, we also examined N400 context effects. We found canonical N400 amplitude effects (Kutas & Hillyard, 1984), consistent with facilitated lexical and semantic processing for expected words. Moreover, in predictive contexts, unexpected words elicited an earlier negativity (about 200-250 ms post-word onset) relative to expected words, strongly suggesting that participants were formulating predictions in our tasks. Again, this pattern was found irrespective of comprehension modality.

Concerning the main goal of the current thesis, we found that electrophysiological activity reliably captures different degrees of context-driven anticipation about upcoming words. The SNPs observed had the following features: (1) they had a graded amplitude modulation across conditions, whereby mean amplitudes were most negative in HC, followed by LC, and NS (with a significant linear tendency across them), (ii) they were most pronounced at left anterior sites, and (iii) they showed a progressive and a sustained increase in time of the difference between HC and the other, less predictive conditions (LC and NS), reaching maximal differences immediately before the presentation of the word (in the 200 ms window before its presentation). With respect to the functional basis of this negativity, Study 2 revealed very similar effects in reading comprehension which, along with the use of the same materials in both studies, lends support to the partake of a common high-level operation in its origination.

Due to their psychophysiological features, SNPs are taken to reflect relatively enhanced neural activity in the cortical areas beneath the electrodes in which they are recorded. Importantly, the left anterior distribution of the effect (HC vs. LC, and HC vs. NS) in both modalities of comprehension argues strongly against an interpretation based on mere sensory expectation for the upcoming stimulus, as this would elicit SNPs with sensory-specific distributions in each input modality (Brunia & Van Boxtel, 2004). Interestingly, resembling pre-stimulus left anterior SNPs have been associated with the retention of verbal items in WM (Khader, Ranganath, Seemüller, & Rösler, 2007; Rolke, Heil, Hennighausen, Häussler, & Rösler, 2000; Ruchkin et al., 1997; Ruchkin, Johnson, Canoune, & Ritter, 1990), as well as with the retrieval of verbal information (Heil, Rösler, & Hennighausen, 1996; Rösler, Heil, & Hennighausen, 1995). A series of studies (reviewed in Rösler et al., 1997) recorded SNPs associated with storage and retrieval of verbal content from long-term memory (**Figure 17**). In these studies, participants had to learn associations between a drawing and either words or positions. In the learning phase, the drawing appeared and, seven seconds later, the associated words or positions were presented. At the beginning of learning, there were no differences during the anticipatory interval, between the drawing and the associated item. But, as participants learnt the associations and therefore could anticipate and retrieve the second item, a SNP developed in the anticipatory interval. Interestingly, the scalp distribution of the SNP was material-specific, such that it was maximal at parietal sites (Pz) for positions, and at left anterior sites (F3) for verbal stimuli, with the latter being very similar to the SNP observed in the current studies. Moreover, the amplitude of the SNPs increased with retrieval difficulty, increasing with the number of items that had to be retrieved from memory, and this increase was also most prominent over the content-specific regions.

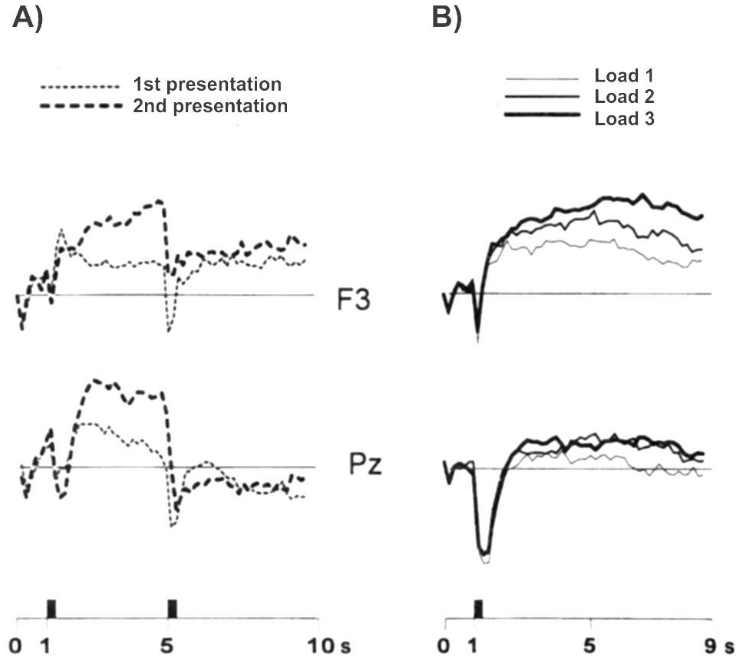

Figure 17. Slow negative potentials elicited in an associative learning task, during the interval between the first stimulus (a drawing) and the associated item (one, two or three words). A) An SNP developed as participants learnt the association between a drawing and one, two or three words (larger amplitude for the 2nd compared to the 1st presentation). The differences were most prominent at verbal-specific sites (at F3, located over language-related regions). **B)** When averaging the trials as a function of the number of words to-be-retrieved (one, two or three), the SNP becomes larger for as the number of associations that must be retrieved increases. *Adapted from Rösler, Heil & Röder (1997).*

The idea that the SNP observed in the current studies is associated with mechanisms that operate during retrieval and/or maintenance of verbal information fits nicely with the idea that, in our studies, participants were either pre-activating information were maintaining pre-activated information active during the anticipatory interval. However, one should remain cautious when drawing connections between our findings and the abovementioned studies, as the tasks used are very different. When this thesis was devised, there was a paucity of studies adopting similar approaches, which difficulted the interpretation of the effect. One of the few comparable ERP studies, Besson and colleagues (1997), contrastingly found more negative amplitudes for LC than HC in written comprehension. A critical difference is that this study employed proverbs as HC sentences, which might involve qualitatively different pre-activation processes (Cacciari, 2014; Cacciari & Tabossi, 1988). Indeed, one study found that, when reading idiomatic expressions such as proverbs, N400 amplitudes between contextually related and unrelated words were indistinguishable, suggesting absent semantic expectancy (Rommers, Dijkstra, & Bastiaansen, 2013). Unlike novel sentence constructions, idiomatic expressions (like proverbs) might be stored in memory as a unitary representation, rendering unnecessary the pre-activation of single words. If this is the case, the finding by Besson and colleagues (1997) would be consistent with ours, as proverbs would yield less semantic expectancy for upcoming words than LC contexts.

In recent years and in parallel to the development of this work other researchers have begun to adopt similar approaches. These studies have consistently uncovered larger negativities before critical words in HC than LC contexts using different materials (Grisoni, Miller, & Pulvermüller, 2017; Li, Zhang, Xia, & Swaab, 2017). For instance, a very similar SNP was described by Li and collaborators (2017) in Mandarin Chinese. They used HC and LC sentences with a critical verb that served to predict (or not) a critical noun appearing two words after ('To practice calligraphy, my brother *bought*_brand-name *brush* pens and took them home') (the critical verb and critical noun are in italics) – in HC contexts, an anterior SNP appeared 480 ms after the onset of the critical verb and lasted for about 900 ms, until the critical word appeared.

Another interesting study was conducted by Grisoni et al., (2017), who presented HC contexts that were strongly predictive of either hand-related ('I take the pen and I write') or face-related verbs ('I take some grapes and I eat') that appeared after 1.5 seconds. To construct LC contexts, they used the same sentences with negation (e.g. 'I take the pen and I do not write'). Interestingly, a SNP developed prior to words in HC (but not in LC contexts) and differences were located over the motor areas related to the expected concept –dorsolateral for hand-related words (e.g. "write") and ventral for face-related words (e.g. "talk"). This pattern of semantic somatotopy led the authors to conclude that the effect was related to the pre-activation of certain semantic features. Although differences in the tasks employed call for caution in adopting a common interpretation for these effects, the convergence among the findings suggest a common underlying cognitive phenomenon. Taking notice of this converging pattern, Pulvermüller and Grisoni, (2020) have recently coined the term *semantic prediction potential* (SPP) to refer to these family of negativities, including the SNPs described in this dissertation.

When interpreting the current findings as being related to semantic prediction, an arguably critical issue is that we do not have direct evidence that participants were retrieving or maintaining specific representations. In this sense, we gained convergent empirical support for this from N400 context effects in the post-word stage. We found the expected N400 modulation as a function of cloze-probability (Kutas & Hillyard, 1984) in both comprehension modalities, which supports facilitated processing for expected words due to their lexical or semantic pre-activation (Federmeier & Kutas, 1999). This interpretation could be disputed considering the coexistence of prediction and integration effects on the amplitude of the N400 (see Section 3.1.3).

However, there is convincing evidence supporting that N400 context effects *also* reflect prediction (Lau et al., 2008), especially earlier portions of the N400 context effect (Nieuwland et al., 2019). Furthermore, we observed that, after predictive contexts, incongruent words elicited a larger negativity than congruent words as early as 200-250 ms after word onset, which is in line with previous studies using predictive discourse contexts (Boudewyn, Long, & Swaab, 2015) or moderately predictive contexts (Brothers et al., 2015). This early negativity strengthens the idea that participants were indeed generating predictions in our studies, perhaps even at the level of word form. Similar N200-N250 effects are found in response to phonological mismatches (van den Brink, Brown, & Hagoort, 2006; Newman & Connolly, 2009) and, in the visual domain, N250 responses are reduced for repeated words (Holcomb & Grainger, 2006), in line with facilitated orthographic processing (Morris, Grainger, & Holcomb, 2008). Hence, the early congruency effect may reflect either the detection of mismatching input (i.e. incongruent words) or a processing head start for expected words.

Well-accepted interpretations of the N400 context effect support that participants in our studies were generating expectations of upcoming words at least at the semantic level. For this reason, we adopted the broader term *semantic anticipation* as an umbrella term to refer to the processes that subserve the elicitation of these signatures.

9.1.2 Transient alpha power decreases

Along with findings in the ERP domain, in **Study 3** we investigated oscillatory activity in pre-word interval by performing time-frequency analyses on the datasets of Study 1 and Study 2. We focused on alpha (8-12 Hz) and beta (13-25 Hz) power decreases, which have been consistently associated with anticipation and prediction in other cognitive domains, both as having a role in optimizing sensory processing (Jensen, Bonnefond, & VanRullen, 2012) and also, in recent years, in boosting the retrieval and encoding of information (Hanslmayr et al., 2012). We found that, prior to words in HC contexts (relative to LC), alpha power oscillatory activity (8-12 Hz) decreased transiently (about –700 to –400 ms before word onset) over frontal areas in both modalities of comprehension. Similarities in spectral, temporal and spatial aspects of the observed alpha power oscillatory decreases suggest the engagement of partially shared mechanisms between modalities.

To what extent might the observed alpha power decreases and SNPs reflect a similar process? On the one hand, both types of neural signatures tend to be found over task-relevant areas and are taken to signal enhanced activation of underlying cortical assemblies (Khader, Jost, Ranganath, & Rösler, 2010; van Boxtel & Böcker, 2004). Interestingly, one study found that alpha power decreased during long-term retrieval over material-specific sites (frontal for objects and parietal for locations) and that these decreases are modulated by the number of to-be-retrieved items, suggesting that it is functionally related to the retrieval of representations (Khader & Rösler, 2011; see also, Burgess & Gruzelier, 2000). Further studies have found that alpha power is decreased during deeper semantic processing (Hanslmayr et al., 2009; Klimesch, 1996, 1999). In fact, the finding of modality-independent, frontal alpha power decreases in HC relative to LC might be more in line with memory-related functions of alpha oscillations. There are recent theories that propose that neural desynchronization at low-frequency bands (i.e. alpha and beta) mediate the representation of information (Hanslmayr et al., 2012) by providing windows of opportunity to enhance cognitive processing capacity (Griffiths et al., 2019).

Overall, both neurophysiological signatures observed in the pre-word interval – left anterior SNPs and alpha power decreases – are consistent with increased processing demands in predictive contexts and might be associated with the need to retrieve and/or maintain information in HC (relative to LC). Yet, differences in their temporal dynamics (sustained vs. transient) and scalp distribution suggest that they do not reflect the same underlying operation. One difference is that oscillatory activity (but not ERPs) unveiled modality-dependent activity as well – in reading comprehension, alpha power decreases extended to occipital sites (associated with visual processing). This modality-dependent alpha power decrease was not observed for the spoken comprehension task, which could be explained by methodological as well as physiological reasons (see Section 7.4). Interestingly, posterior transient alpha power decreases have been functionally linked with improved visual processing of upcoming stimulus (van Dijk, Schoffelen, Oostenveld, & Jensen, 2008; Jensen & Mazaheri, 2010; Mathewson et al., 2014), which might agree with the beforementioned early congruency effects (about 200 ms post-word) for final words.

9.1.3 Sustained slow negative potentials as online indices of prediction over the course of sentence comprehension

The second study showed that the differences captured in the pre-word interval begun earlier in time and developed gradually over the course of sentence processing. In HC sentences (relative to LC and NS), a sustained negativity appeared progressively at frontal and central sites. In the domain of sentence comprehension, frontally-distributed sentence-level negativities have been consistently associated in the past with increased WM demands when processing sentences with syntactically complex structures (Fiebach et al., 2002; King & Kutas, 1995; Müller et al., 1997), which has been mostly attributed to the active maintenance of information (Clifton & Frazier, 1989; Gibson, 1998). By extension, we proposed that the observed relative sentence-level negativity in predictive sentence contexts might similarly reflect enhanced WM demands incurred by maintaining pre-activated information (see Section 6.4).

But what is the relationship, if any, between prediction and WM? Lau and colleagues (2013) (see also Kamide, 2008) highlighted that prediction does not simply involve pre-activating representations, as this could be similarly achieved through automatic and passive spread of activation (i.e. priming). Prediction involves an additional step of 'committing' to the pre-activated information by adding it into the ongoing and active context representation, in their words, *"In distinguishing between these two mechanisms [passive spread of activation and prediction], we appeal to the existence of some form of working memory or 'focus of attention' that holds the contextual representation [...] prediction refers specifically to mechanisms by which the contextual representation, held in WM, is updated in advance of the input"*. This view implies that pre-activated representations are incorporated online into the ongoing context representation as processing continues, rather than remaining 'unintegrated' until they can be used, as the theories proposed to explain anterior negativities in the context of syntactic parsing suggest (Clifton & Frazier, 1989; Gibson, 1998). For pre-activated representations to be of any use to guide comprehension, they must lead to a new context representation that conveys the message-level information necessary to further pin down contextually appropriate words. But insofar as constructing and incrementally updating a context representation is a normal part of comprehension, it is difficult to fathom why pre-updating the ongoing context representation should be particularly taxing for WM by itself, at least to the point of eliciting detectable differences in neural responses in adept, young adult comprehenders (as in Study 1 and Study 2).

As explained in the General Introduction, prediction in sentence comprehension can be upregulated and downregulated by goals and strategies (see Section 3.1.4). An interesting parallel is found in discourse comprehension, in which some discourse-based inferences are extracted routinely and online as the sentence unfolds, whereas others are formed only under certain circumstances, if strategically viable and advantageous (Graesser, Singer & Trabasso, 1994), such as when the context is highly constraining or the interval between the context and final word is of at least 750 ms or more (Calvo & Castillo, 1996; Calvo, Castillo, & Estevez, 1999). Critically, some theoretical models posit that these inferences are drawn from and used to 'pre-update' a *situation model* that readers actively build during comprehension, a mental representation of the events, actions, people, and the situation the text is about that is constructed in parallel to the obligatory *text* representation (Johnson-Laird, 1980; van Dijk & Kintsch, 1982; Zacks, Mar, & Calarco, 2008). Similarly, the observed sentence-level negativity in predictive sentence contexts could be a proxy indicating that readers were actively constructing a rich context representation and holding it in

WM to formulate predictions and guide processing. In this line, Ferreira & Chantavarin (2018) recently hypothesized that the construction of a rich context representation in language comprehension induces a forward-looking state of 'preparedness'.

Speculatively, we suggest that this notion of a forward-looking state intriguingly connects with a well-established idea in theories of flexible cognitive control (Braver, Paxton, Locke, & Barch, 2009), which distinguishes between two modes of cognitive control, 'proactive' and 'reactive' control mode. The proactive control mode is seen as a form of 'early selection', whereby the active maintenance of goal-relevant information in WM in an anticipatory and sustained manner biases future behaviour, including attention, perception or motor systems. In turn, the reactive mode consists of a more laissez-faire processing mode, in which attentional control acts as a corrective mechanism only when interferent events occur. Most interesting, proactive control should be characterized by anticipatory and sustained activity on the lateral prefrontal cortex (PFC) activity reflecting the maintenance of task goals (Braver et al., 2009; Miller & Cohen, 2001; Paxton, Barch, Racine, & Braver, 2008), which fits well with the sustained and frontally-distributed nature of the scalp negativities described in the current studies. On the other hand, reactive control would be reflected in transient activity of the lateral PFC, and in the recruitment of regions linked to conflict monitoring, such as the anterior cingulate cortex, and with posterior regions involved in associative connections. Note that this distinction between proactive and reactive control is reminiscent of that of 'predictive' and 'incremental' processing in language comprehension. However, it does not see each mode as mutually exclusive but relies on cognitive control to flexibly balance between them (Barch & Ceaser, 2012).

Drawing upon these intuitions, in a new study (in preparation) we continued to investigate the proposed forward-looking quality of the sentence-level negativities elicited in our experimental paradigm. We reasoned that, if these signatures are associated with prediction, they would not emerge when future input is known. To test this, we manipulated whether participants could know the final word of the sentence (i.e., 100% cloze-probability). Before reading a sentence, a single word was presented (cue) that could or could not correspond to the final word of the next sentence. For example, the cue 'couch' would be informative of the next sentence 'he is taking a nap on the <u>couch</u>', but the cue 'drawer' would not be informative in this case. Critically, participants discovered whether the cue was informative or not by virtue of a semantically-related noun in the sentence that appeared either at the 3rd position or 5th word position (e.g., 'nap' in 'he is taking a <u>nap</u> on the couch'). Therefore, we controlled *when* key information to resolve uncertainty about the final word was delivered, either earlier or later in time. As in our previous studies, we inserted a 1 second break between the sentence context and the final word. Our hypothesis was that sentence-level negativities would emerge when the cue was uninformative, and uncertainty about the upcoming word remains. In turn, when uncertainty about the final is completely resolved thanks to the cue, the negativity would cease to emerge. In line with this, a preliminary analysis (N = 20) revealed that a larger negativity emerged as soon as the cue turned out to be uninformative, compared to when it was not. These findings are consistent with readers switching to the proposed proactive control mode *when* they need to figure out what the final word is, but not when it is already known.

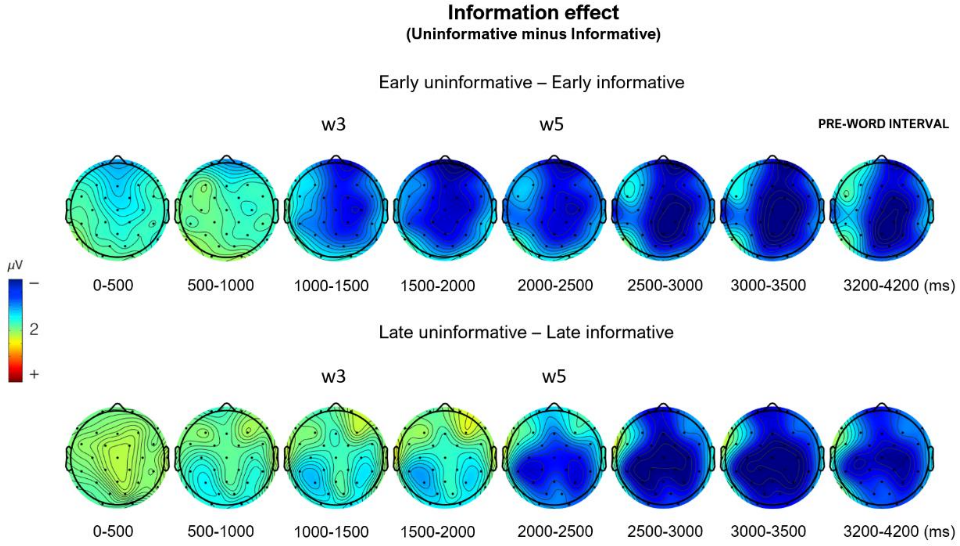

Figure 18. Scalp maps displaying the temporal evolution of the sentence-level negativity after encountering critical words earlier (w3) or later (w5) during sentence processing. The scalp maps are based on mean amplitude differences (uninformative minus informative) in eight consecutive epochs (seven context words and the pre-word interval). A centro-parietal sustained negativity emerged when comprehenders found out that the cue that was initially provided (before the sentence onset) was uninformative about the sentence-final word (after reading 'nap' in the sentence 'He is taking a nap on the couch' when the initial cue was 'drawer'), compared to when the cue was informative ('couch'). The critical word ('nap') appeared at the third (top row) or fifth (botton row) position in the sentence. The sustained negativity was sustained until the presentation of the sentence-final word ('couch'). The ERPs are low pass filtered at 5Hz. The range of voltage values for the maps is –2 µV.

Note that the emergence of the sustained negativity cannot be explained by the sum of transient word-level effects (King & Kutas, 1996). As expected, critical words in the uninformative condition elicited a larger N400 response, at the 3rd or 5th word, but then the difference in the overall amplitude of the negativity was equalized between both uninformative conditions. This suggests that the sentence-level negativity reflects an independent process, on which transient, word-level responses are overlaid. Furthermore, the sentences were the same in informative and uninformative conditions, and thus the effects in the emergence of the negativity cannot be attributed to differences the syntactic structure.

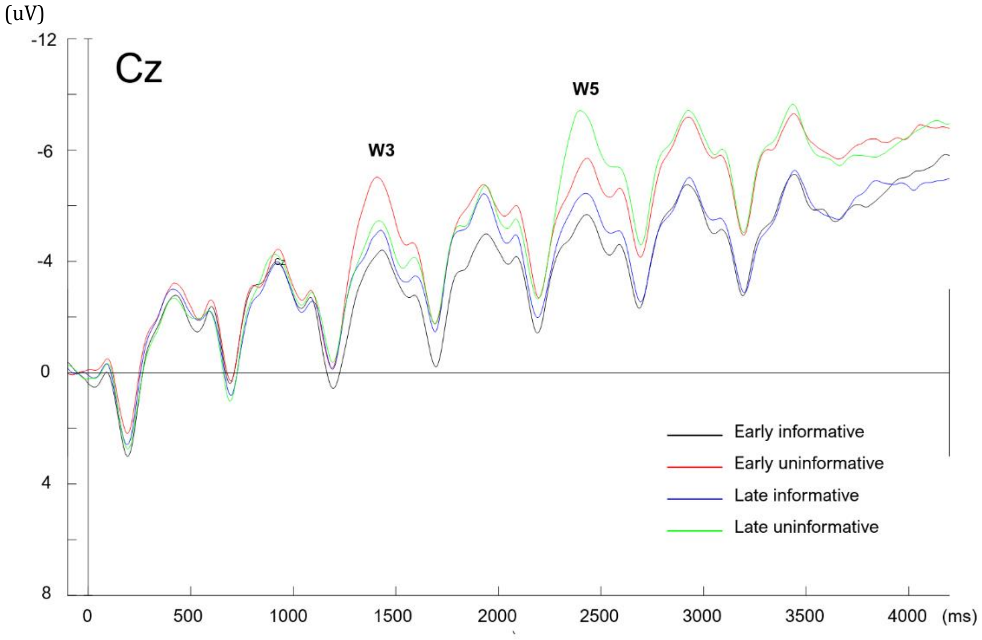

Figure 19. Sustained negativities after encountering critical words earlier (w3) or later (w5) during sentence processing. Grand-averaged ERPs during the sentence context at the Cz electrode for each condition. A larger N400 response was elicited by uninformative compared to informative conditions at the position when the critical noun was found, either early (3rd word; w3) or late (5th word, w5). From then onwards, uninformative conditions continued to exhibit a larger sustained negativity until the presentation of the final word (compared to informative conditions). The data is low pass filtered at 5Hz.

In sum, the amplitude of the sentence-level negativity was reduced when readers could be certain about the sentence-final word. Curiously, this finding aligns well with the ‘information-extraction hypothesis’ proposed by Morís et al. (2013) to explain learning-induced amplitude changes in the stimulus-preceding negativity (SPN) in reinforcement learning tasks (see also Ren et al., 2017). As soon as feedback, which can be paralleled with the sentence-final word in our studies, did not provide information gain because the association was learnt, the SPN was greatly reduced, compared to earlier stages of learning in which the feedback was still informative. In our new experiment, the sustained negativity emerged as soon as the cue turned out to be uninformative, and therefore there was still relevant information to extract from the final word in order to achieve comprehension.

At first, these findings might appear inconsistent with the results in Study 1 and Study 2. If the sentence-level negativity is associated with the information gain of the upcoming input, then one would expect to find more negative amplitudes for LC than HC sentence contexts (the opposite of what we found), given that there is more information to gain from final words after less predictive LC sentence contexts. However, one way to parsimoniously reconcile these findings is through the idea that, in language comprehension, the construction of a rich context representation is a prerequisite to induce a state of preparedness in which expectations are generated (Ferreira & Chavantarin, 2018). In fact, in Study 2, we did not find differences between the sentence-level negativities in LC and NS contexts, which is consistent with the underlying process being dependent upon having *enough* information to build a rich context representation. This view is also consistent with the information-via-desynchronization hypothesis (Hanslmayr et al., 2012) whereby the modality-independent relative alpha power decreases in HC sentence contexts in Study 3 might represent an information-rich scenario in which more specific information is available to construct a context representation, perhaps with an episodic quality as in the case of situation models in discourse comprehension (van Dijk & Kintsch, 1982). We proposed that this is not only due to the incremental construction of a representation of meaning, but that, based on the anticipatory SNPs in the ERP domain, the observed alpha power decreases also have a forward-looking quality. Interestingly, (Klimesch, 2012) proposed that alpha oscillations reflect ‘semantic orientation’, which refers to the ability to be consciously oriented in relation to the external world. This emphasizes a forward-looking connotation as it relies on the ability to selectively orient attention and access stored knowledge, such as language.

This proactive control mode may not be habitual in everyday language use, but there might be certain situations that could call for and benefit from more invested comprehension. For example, when learning a new language as an adult, it might be beneficial to invest resources in forming elaborate context representations via prediction to effectively orient behaviour, such as attention, to future information that will be most critical for learning. Another situation in which the postulated proactive control could be at play is during simultaneous language interpretation, in which actively sustaining an enriched sentence or discourse representation could critically help catch up with the interlocutor’s pace and even allow to modulate prosody accordingly in a fast and online manner. The electrophysiological signatures described in the current dissertation could serve as reliable markers of the engagement of this mode of processing, an idea that promises an exciting path for future research to understand when and how prediction supports different endeavours in language comprehension.

9.2 Predictive language processing in Parkinson's disease

Finally, in **Study 4**, we adapted the reading comprehension task for its application with individuals with Parkinson's disease (PD) within a grand-scale longitudinal study lead by the Movement and Disorder Unit in Hospital Sant Pau i de la Creu. Cognitive deficits are common in PD, with some PD patients meeting criteria for mild cognitive impairment (PD-MCI). An unaddressed question was whether these cognitive deficits may impair language prediction in this clinical population, since predictive language processing seems to be less in other populations with reduced cognitive abilities like aging individuals (e.g., Federmeier, McLennan, De Ochoa-Dewald & Kutas, 2002). We reasoned that similar changes in the ability to formulate linguistic predictions could account for the difficulties that some PD patients face with language comprehension, especially as the disease progresses (Pell & Monetta, 2008).

We found that PD patients with normal cognition (PD-NC), these patients exhibited the expected predictive language processing ERP effects (pre- and post-word) (relative to matched healthy controls). This indicated that they could not only successfully construct a sentence representation but also pre-activate certain aspects of future linguistic input to prepare ahead of time. It should nonetheless be stressed that all patients were tested during their stable doses of dopaminergic compensation, so it cannot be discarded that subtler problems could be evidenced off medication. This is plausible considering that dopamine may mediate semantic activation (Kischka et al., 1996) and there is evidence of a delayed time-course of lexical and semantic activation in PD (off and on medication) (Grossman et al., 2002; Angwin et al., 2007; Angwin et al. 2009; Copland et al., 2000; Copland et al. 2003).

We concurrently evaluated the impact of cognitive decline on predictive language processing in PD. We found that PD-MCI is associated with normative N400 amplitude differences, but with a prolonged N400 congruency effect indicating difficulties dealing with input that disconfirms strong predictions. Due to its latency, the prolongation may be attributed to integration difficulties (Nieuwland et al., 2020; see also, Lau et al., 2016). In PD and in other subcortical neurodegenerative diseases, frontal-executive alterations are prototypical due to the deterioration of fronto-striatal projections (e.g., Aarsland et al., 2010). This may involve damage to IFG regions, which would be consistent with sentence integration problems (Hagoort, Baggio & Willems, 2009) and difficulties processing semantic anomalies (Baumgaertner, Weiller & Büchel, 2002; Kiehl et al., 2002; Newman et al., 2001). On the other hand, the design of the current task cannot rule out that the effect could reflect problems with lexical and semantic activation. In this respect, damage in anterior and posterior IFG may also lead to difficulties in using top-down information to mediate controlled retrieval and selection of representations (Lau et al., 2008) (see Section 3.1.3.2). Thus, an alternative explanation could be that PD-MCI patients might have difficulties retrieving the contextually appropriate word while/or inhibiting an interfering incongruent word. Furthermore, in PD-MCI, temporo-parietal structures start to deteriorate, and recent clinical studies indicate that the emergence of posterior-cortical deficits involving language, memory or visuospatial alterations might serve as early cognitive markers of the progression to PD with dementia (PDD) (Pagonabarraga & Kulisevsky, 2012; Horta-Barba et al., 2020). In support of the involvement of temporal circuits in the elicitation of the N400 congruency effect, we found that it correlated with verbal fluency in all PD patients.

What do the results in Study 4 tell us about the SNP? Perhaps the most interesting result is that the SNP was diminished in PD patients with lower verbal fluency scores. Interestingly, the anatomical substrates of verbal fluency scores (both semantic and phonological) have been localized in the left IFG and insula (Matthijs et al., 2016). Considering its left anterior scalp distribution, the left IFG could be a potential cortical generator of the pre-word SNP. Damage to left IFG areas would thus parsimoniously account for the observed pattern of correlations between verbal fluency scores and ERP effects, impacting those more reliant on frontal function (SNP and N400 congruency effects) while leaving temporal-dependent effects (N400 constraint effects) relatively spared. At the functional level, this would involve difficulties implementing top-down predictions to select and retrieve expected representations (Lau et al., 2008).

Furthermore, as has been explained before, the implication of the left IFG and, more generally, left PFC regions in the elicitation of the SNP would be in line with the idea that it is associated with the active maintenance of information, which is a marker of proactive control (Braver et al., 2009). Intriguingly, the dopaminergic system has a pivotal role in balancing between proactive and reactive processing modes (Braver, 2012). Of special interest here, proactive control relies on temporally precise fronto-striatal dopamine interactions, in particular, phasic DA signals maintain inputs in the prefrontal cortex (PFC). Disrupted DA signals would lead to only transient activation in the PFC, preventing the active maintenance of information in WM (Braver & Cohen, 2000). As a result, proactive control may be disrupted in clinical populations with dysfunctional prefrontal and DA function, schizophrenia, older adults, and speculatively also PD patients, who might rely more on reactive control (Barch & Ceaser, 2012). Note that reliance on reactive control due to poor DA function would account for the pattern of ERP effects associated with lower verbal fluency scores in PD – N400 responses would be prolonged due to the need to process information as it flows in without the advantage of anticipation.

In sum, this study provided at least three crucial insights. First, that predictive language processing (at least as evaluated with the current task) is preserved in PD patients on dopaminergic compensation, but that further cognitive constraints imposed by MCI leads to difficulties in dealing with words that disconfirm predictions. Second, from a clinical lens, it supports that PD-MCI patients may present cortical deterioration in other networks beyond frontal circuits, probably also encompassing temporal-dependent areas. And, mechanistically, it suggests that the left IFG may be involved in the emergence of the pre-word SNP, which agrees with its prominently left frontal lateralization. Future studies with different neuroimaging techniques could elucidate the structural underpinnings of the captured processes, both in healthy participants and PD patients.

Finally, identifying early cognitive indicators of the progression from PD-MCI to PDD is a pivotal challenge in clinical research of PD. The ability to predict the prognosis of cognitive deterioration in these patients will contribute to the improvement of the preventive and palliative treatments of the disease. Most of the scientific effort has been directed at ameliorating their motor deficits, as they are the most compelling, but cognitive and language-related problems may be as distressing and invalidating in the life of PD patients. To advance in this important issue, the Movement Disorders Unit of Hospital de Sant Pau i de la Santa Creu launched a grand-scale longitudinal study that will help pinpoint early cognitive markers of cognitive deterioration in PD. The task in Study 4 was included in the experimental protocol of this ambitious project, and the same participants that participated in Study 4 completed the same experiment one year later. The current findings have importantly helped to isolate the N400 congruency effect as a potential predictor of the transition from PD-MCI to PDD. We will thus next prospectively track within-subject changes in the N400 congruency effect and explore changes in other measures, with a special interest in detecting differences in cognitive processes that require the participation of temporo-parietal dependent circuitry given its importance in PD-MCI (Horta-Barba et al., 2020).

General Conclusions and Limitations

10 General Conclusions and Limitations

In order to shed light on the neural mechanisms subserving the generation of semantic predictions during sentence comprehension, we developed a new approach that allowed to explore anticipatory activity associated with the semantic prediction of words during sentence comprehension.

With respect to the main goal of the current thesis, we investigated electrophysiological activity associated with semantic anticipation during spoken and reading comprehension in healthy, young adult population (Study 1, 2 and 3). In both modalities of comprehension, we found that slowly accruing, sustained negative potentials appear between a sentence context and its final word, with a larger amplitude when the context is strongly constraining and thus predictive of that word. Furthermore, tracking slow activity (below 5 Hz) during word-by-word reading comprehension (Study 2) revealed that the pattern of context-related amplitude differences emerged earlier and progressively over the course of sentences, in the form of fronto-central, sustained negativities. Concurrently, time-frequency analyses (Study 3) revealed that transient, relative alpha power decreases also preceded words in strongly constraining contexts in both modalities of comprehension. In that case, modality-specific preparation was observed in reading, but not in spoken comprehension, being the effects overall weaker in the latter modality. Taken together, these results are in line with the idea that, when the sentence context permits it, readers and listeners mobilize cognitive resources online to prepare for upcoming words ahead of time, and that this can be monitored with EEG. Specifically, the spatial, temporal, and spectral features of the observed signatures, along with their modality-generality, match with well-established effects associated with mechanisms that support and boost retrieval and/or storage of representations. In agreement with this, post-word congruency effects were detected as early as 200-250 ms after word onset, in line with either early mismatch detection or word processing facilitation by virtue of prior pre-activation of probable features.

Finally, in Study 4, we adopted the same approach to examine the status of mechanisms associated with semantic prediction in a large sample of individuals with PD and with PD and MCI on dopaminergic compensation. We showed that PD patients exhibited preserved signatures of semantic anticipation and appropriate semantic processing in reading comprehension (relative to matched controls). Besides, additional cognitive constraints imposed by PD-MCI led to prolonged N400 congruency effects after predictive contexts, pointing to difficulties in dealing with contextually unexpected words (relative to PD patients without MCI matched in age, gender, and years of education). From a clinical standpoint this importantly suggests that – on top of deterioration in frontal-dependent circuits – temporal-dependent circuits are affected in PD-MCI, and pinpoint N400 congruency effects as a potential marker of the advent of problems with semantic processing in this clinical population. Lastly, in all PD individuals, lower verbal fluency scores were linked with altered SNPs and prolonged N400 congruency effects. We concluded that cognitive deficits in PD spare semantic anticipation and processing, but the deterioration of frontal as well as temporal-dependent mechanisms lead to difficulties with utilizing sentence contexts to prepare for upcoming words and with later slower, less efficient semantic processing.

Overall, this body of work demonstrates that electrophysiological signatures can reliable capture how contextual predictiveness impacts neural processing at different time-scales during sentence comprehension, strengthening the idea that, when the sentence context enables it, readers and listeners recruit cognitive resources online to effectively prepare for upcoming words. These findings will hopefully pave a path for future basic and clinical research to improve our understanding of what, how, and when readers and listeners engage in predictive language processing during comprehension.

Limitations and future directions

Despite the empirical contribution of the current findings, there are some limitations that must be considered to understand the scope of these findings and to inform future research directions:

- It is likely that the sentence contexts that we used in these studies do not differ only in terms of predictiveness. Sentence contexts that are less constraining tend to contain terms with more vague meanings than highly constraining contexts. This might lead to differences between conditions in the level of attention or arousal deployed during comprehension, without the need to call for prediction. This is a difficult problem to avoid in sentence comprehension tasks because predictiveness of the sentence context is intrinsically associated with the degree of information that it provides. However, as we have explained in the General Discussion, we consider that, in language comprehension, prediction might go hand-in-hand with the construction of a rich context representation (see also Ferreira and Lowder, 2018), and therefore, in this view, one might wonder to what extent factoring out differences in attention, arousal or motivation makes sense when approaching predictive language processing. Future studies are needed to carefully manipulate these variables and understand how they might impact the elicitation of the observed effects.
- We do not have direct evidence that participants were indeed pre-activating certain representations in our tasks. Based on abundant prior research, the N400 amplitude

modulations that we observed index prediction-driven lexical-semantic facilitation (Lau et al., 2008). However, as we explained in the General Introduction, these N400 context effects could also be in part attributed to integration. Some authors have opted to explicitly instruct participants to predict and collect their responses (Brothers et al., 2015), but this might introduce additional decision-related processes. To obtain stronger evidence for prediction without introducing confounding variables, one possibility could be to employ materials in which the pre-final word can already elicit prediction-inconsistent responses (as in Section 3.1.3.4).

- An important feature of the employed materials is that the final words were incongruent in 50% of the trials. This is common practice in tasks that address linguistic prediction because it allows examining the effect of prediction mismatch (e.g., Besson et al., 1997; Grisoni et al., 2017; Li et al., 2017; Rommers et al., 2017; Wang et al., 2017). Yet, predictive validity – the rate to which predictions are successful in a task or experimental block – influences the degree to which individuals engage in prediction (Brothers et al., 2017; Lau et al., 2013). It will be important that future studies investigate how this feature of the design might have influenced in the elicitation of the SNPs and alpha power decreases, a question that we have begun to address.

- We did not consider the influence of the syntactic structure in the evolution of sentence-level effects (Study 2). It is possible that differences take off after processing certain types of words. In fact, if the negativity unfurls due to pre-activation and/or pre-updating, it is likely that it is contingent on words that provide more information, like the main verb or other content words. For example, Li et al (2017) found that the SNP arose as soon as participants read the verb that provided crucial information to predict an upcoming word. Future studies could systematically manipulate the syntactic configuration of sentences to research whether certain words are more critical in the elicitation of these negativities.

- A question that remains unanswered is to what extent the temporal predictability of the final word played an important role in eliciting the observed effects, given that temporal predictability is an important factor in sensory prediction and anticipation (Arnal & Giraud, 2012). In this respect, one previous study suggested that mechanisms supporting semantic prediction do not depend on precise temporal prediction (Lau & Nguyen, 2015), Future research could help clarify to what extent the temporal predictability of critical words might be a necessary condition for the elicitation of the observed sustained potentials.

- The way in which language is presented in our tasks is evidently far from how we face language in real life. While sentences may occasionally get interrupted (i.e. the speaker loses her train of thought), they are not interrupted systematically and expectedly. Further, in natural language, sentences do not tend to be as predictive as the ones used in the highly constraining condition in our tasks (Luke & Christianson, 2016). Among other factors, this clearly limits the extrapolation of the current findings to more naturalistic language situations, which is a trade-off that is intrinsic to most experimental approaches. Yet, as we have previously discussed, studying extreme situations, in which prediction is highly encouraged, can provide useful insights about how and when individuals might strategically engage in some forms of processing. These approaches allow obtaining neural signatures that could be then used by future studies to improve our understanding of how we process language in other, more naturalistic situations.

- We might have missed more fine-grained differences in the ERP responses between groups in Study 4. Our aim here was to look only at mean amplitude differences, however, a number of previous studies using analytical approaches have reported altered onset (Kutas et al., 2013; Angwin et al., 2012) or peak latencies (Kutas et al., 2013) of N400 context effects in PD. It would be interesting to reanalyze the data computing these parameters as well to find out if there are subtler group differences that cannot be captured with cluster-based permutation analyses. A fine-grained approach would also be more appropriate for longitudinal studies looking for neurophysiological markers of the transition from PD to PD-MCI, or PD-MCI to PDD, in which subtle changes in the N400 response would be of utmost informative value.

- In Study 4, we did not find significant differences in the SNP between PD-MCI or in their control group. However, visual inspection of the ERP waveforms showed the expected tendency in both groups and pointed to a greater reduction of the effect in the PD-MCI group. It is possible that we had insufficient power to detect differences in these groups due to few trials or subjects. This is also plausible considering that working with patient data leads to a larger loss of trials during pre-processing. Future studies that aim to study SNP with similar clinical populations could benefit from using a larger amount of trials (more than 40 trials per condition) or, alternatively, increase the sample size (more than 20 participants per group).

References

11 References

Aarsland, D., Bronnick, K., Williams-Gray, C., Weintraub, D., Marder, K., Kulisevsky, J., et al., (2010). Mild cognitive impairment in Parkinson disease: a multicenter pooled analysis. *Neurology* 75, 1062–1069. Http:// dx.doi.org/10.1212/WNL.0b013e3181f39d0e.

Ahveninen, J., Huang, S., Belliveau, J. W., Chang, W. T., & Hämäläinen, M. (2013). Dynamic oscillatory processes governing cued orienting and allocation of auditory attention. *Journal of Cognitive Neuroscience*, 25(11), 1926–1943. https://doi.org/10.1162/jocn_a_00452

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30(3), 191-238. doi: 10.1016/0010-0277(88)90020-0

Angwin, A. J., Arnott, W. L., Copland, D. A., Haire, M. P. L., Murdoch, B. E., Silburn, P. A., et al. (2009). Semantic activation in Parkinson's disease patients on and off levodopa. *Cortex*, 45, 950–959. Https://doi.org/10.1016/j.cortex.2009.02.012

Angwin, A. J., Chenery, H. J., Copland, D. A., Arnott, W. L., Murdoch, B. E., & Silburn, P. A. (2004a). Dopamine and semantic activation: an investigation of masked direct and indirect priming. *Journal of the International Neuropsychological Society*, 10(1), 15-25.

Angwin, A. J., Dissanayaka, N. N., Moorcroft, A., McMahon, K. L., Silburn, P. A., & Copland, D. A. (2017). A neurophysiological study of semantic processing in Parkinson's disease. *Journal of the International Neuropsychological Society*, 23(1), 78-89.

Angwin, A. J., H. J. Chenery, D. A. Copland, B. E. Murdoch, and P. A. Silburn (2004b). The time course of semantic activation in Parkinson's disease. *Brain and Language*, 91.145–6.

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in cognitive sciences*, 16(7), 390-398.

Arnott, W. L., Chenery, H. J., Angwin, A. J., Murdoch, B. E., Silburn, P. A., & Copland, D. A. (2010). Decreased semantic competitive inhibition in Parkinson's disease: Evidence from an investigation of word search performance. *International journal of speech-language pathology*, 12(5), 437-445.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, 225(2), 82–90. https://doi.org/10.1038/scientificamerican0871-82

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.) (Vol. 8, pp. 47–89). Academic Press. https://doi.org/https://doi.org/10.1016/S0079-7421(08)60452-1

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes*, 26(9), 1338–1367. https://doi.org/10.1080/01690965.2010.542671

Baldo, J. V., Schwartz, S., Wilkins, D., & Dronkers, N. F. (2006). Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. *Journal of the International Neuropsychological Society*, 12(6), 896-900.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, 17(3), 364–390. https://doi.org/10.1016/0010-0285(85)90013-1

Banerjee, S., Snyder, A. C., Molholm, S., & Foxe, J. J. (2011). Oscillatory alpha-band mechanisms and the deployment of spatial attention to anticipated auditory and visual targetlocations: Supramodal or sensory-specific control mechanisms? *Journal of Neuroscience*, 31(27), 9923–9932. https://doi.org/10.1523/JNEUROSCI.4660-10.2011

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in cognitive sciences*, 11(7), 280-289. doi: 10.1016/j.tics.2007.05.005

Bar, M. (2009). The proactive brain: Memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1235–1243. https://doi.org/10.1098/rstb.2008.0310

Barch, D., & Ceaser, A. (2012). Cognition in schizophrenia: core psychological and neural mechanisms. *Trends in Cognitive Sciences*. https://www.sciencedirect.com/science/article/pii/S1364661311002488

Barrett, S. E., & Rugg, M. D. (1989). Event-related potentials and the semantic matching of faces. *Neuropsychologia*, 27(7), 913–922. https://doi.org/10.1016/0028-3932(89)90067-5

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, 14(2), 201–212. https://doi.org/10.1016/0278-2626(90)90029-N

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond ERPs: Oscillatory Neuronal Dynamics. In The Oxford Handbook of Event-Related Potential Components. *Oxford University Press*. https://doi.org/10.1093/oxfordhb/9780195374148.013.0024

Baumgaertner, A., Weiller, C., & Büchel, C. (2002). Event-related fMRI reveals cortical sites involved in contextual sentence integration. *NeuroImage*, 16(3 I), 736–745. https://doi.org/10.1006/nimg.2002.1134

Besson, M., Faita, F., Czternasty, C., & Kutas, M. (1997). What's in a pause: Event-related potential analysis of temporal disruptions in written and spoken sentences. *Biological Psychology*, 46(1), 3–23. https://doi.org/10.1016/S0301-0511(96)05215-5

Bidet-Caulet, A., Barbe, P. G., Roux, S., Viswanath, H., Barthélémy, C., Bruneau, N., Knight, R. T., & Bonnet-Brilhault, F. (2012). Dynamics of anticipatory mechanisms during predictive context processing. *European Journal of Neuroscience*, 36(7), 2996-3004.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Birbaumer, N., Elbert, T., Canavan, A. G. M., & Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiological Reviews*. https://doi.org/10.1152/physrev.1990.70.1.1

Böcker, K. B. E., Baas, J. M. P., Kenemans, J. L., & Verbaten, M. N. (2001). Stimulus-preceding negativity induced by fear: a manifestation of affective anticipation. *International Journal of Psychophysiology*, Vol. 43, pp. 77–90. https://doi.org/10.1016/S0167-8760(01)00180-5

Boudewyn, M. A., Long, D. L., & Swaab, T. Y. (2015). Graded expectations: Predictive processing and the adjustment of expectations during spoken language comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, 15(3), 607-624. https://doi.org/10.3758/s13415-015-0340-0

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.12.010

Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *National Academy of Sciences*.

Braver, T. S., & Cohen, J. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. *Control Cognitive Processes*.

Brink, V. Den, Brown, C. M., & Hagoort, P. (2006). The Cascaded Nature of Lexical Selection and Integration in Auditory Sentence Processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 32(3), 364–372. https://doi.org/10.1037/0278-7393.32.3.364

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2015). Effects of prediction and contextual support on lexical processing: Prediction takes precedence. *Cognition*, 136, 135–149. https://doi.org/10.1016/j.cognition.2014.10.017

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, 93, 203–216. https://doi.org/10.1016/j.jml.2016.10.002

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5(1), 34–44. https://doi.org/10.1162/jocn.1993.5.1.34

Brunia, C. H. M., & Van Boxtel, G. J. M. (2001). Wait and see. *International Journal of Psychophysiology*, 43(1), 59-75. doi: 10.1016/S0167-8760(01)00179-9

Brunia, C. H. M., Hackley, S. A., van Boxtel, G. J., Kotani, Y., & Ohgami, Y. (2011). Waiting to perceive: reward or punishment? *Clinical Neurophysiology*, 122(5), 858-868. doi: 10.1016/j.clinph.2010.12.039

Brunia, C. H. M., Van Boxtel, G. M., & Böcker, K. B. (2011). Negative slow waves as indices of anticipation: the Bereitschaftspotential, the contingent negative variation, and the stimulus-preceding negativity. In The Oxford handbook of event-related potential components, 189-207.doi: 10.1093/oxfordhb/9780195374148.013.0108

Brunia, C. H.M., & Damen, E. J. P. (1988). Distribution of slow brain potentials related to motor preparation and stimulus anticipation in a time estimation task. *Electroencephalography and Clinical Neurophysiology*, 69(3), 234–243. https://doi.org/10.1016/0013-4694(88)90132-0

Brunia, C. H.M., & Van Boxtel, G. J. M. (2004). Anticipatory attention to verbal and non-verbal stimuli is reflected in a modality-specific SPN. *Experimental Brain Research*, 156(2), 231–239. https://doi.org/10.1007/s00221-003-1780-2

Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE transactions on medical imaging*, 18(1), 32-42. https://doi.org/10.1109/42.750253

Burgess, A. P., & Gruzelier, J. H. (2000). Short duration power changes in the EEG during recognition memory for words and faces. *Psychophysiology*, 37(5), 596–606. https://doi.org/10.1111/1469-8986.3750596

Buzsáki, G., & Watson, B. O. (2012). Brain rhythms and neural syntax: Implications for efficient coding of cognitive content and neuropsychiatric disease. *Dialogues in Clinical Neuroscience*, 14(4), 345–367. https://doi.org/10.31887/dcns.2012.14.4/gbuzsaki

Cacciari, C. (2014). Processing multiword idiomatic strings. The Mental Lexicon, 9(2), 267–293. https://doi.org/10.1075/ml.9.2.05cac

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27(6), 668–683. https://doi.org/10.1016/0749-596X(88)90014-9

Calvo, M. G., & Castillo, M. D. (1996). Predictive inferences occur on-line, but with delay: Convergence of naming and reading times. *Discourse Processes*, 22(1), 57–78. https://doi.org/10.1080/01638539609544966

Calvo, M. G., Castillo, M. D., & Estevez, A. (1999). On-line predictive inferences in reading: Processing time during versus after the priming context. *Memory and Cognition*, 27(5), 834–843. https://doi.org/10.3758/BF03198536

Catena, A., Perales, J. C., Megías, A., Cándido, A., Jara, E., & Maldonado, A. (2012). The Brain Network of Expectancy and Uncertainty Processing. *PLoS ONE*, 7(7), e40252. https://doi.org/10.1371/journal.pone.0040252

Cermolacce, M., Scannella, S., Faugère, M., Vion-Dury, J., & Besson, M. (2014). "All that glitters is not… alone". Congruity effects in highly and less predictable sentence contexts. *Neurophysiologie Clinique*, 44(2), 189-201. doi: 10.1016/j.neucli.2014.04.001

Christiansen, M. H., & Chater, N. (2015). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*. Cambridge University Press. https://doi.org/10.1017/S0140525X1500031X

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

Clifton, C., & Frazier, L. (1989). Comprehending Sentences with Long-Distance Dependencies (pp. 273–317). https://doi.org/10.1007/978-94-009-2729-2_8

Clos, M., Langner, R., Meyer, M., Oechslin, M. S., Zilles, K., &Eickhoff, S. B. (2014). Effects of prior information on decoding degraded speech: an fMRI study. *Human Brain Mapping*, 35(1), 61-74. doi: 10.1002/hbm.22151

Cohn, N. (2012). Comics, Linguistics, and Visual Language: The Past and Future of a Field. In Linguistics and the Study of Comics (pp. 92–118). *Palgrave Macmillan UK*. https://doi.org/10.1057/9781137004109_5

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Connor, W. H., & Lang, P. J. (1969). Cortical slow-wave and cardiac rate responses in stimulus orientation and reaction time conditions. *Journal of Experimental Psychology*, 82(2), 310. doi: 10.1037/h0028181

Copland, D. A., Sefe, G., Ashley, J., Hudson, C., & Chenery, H. J. (2009). Impaired semantic inhibition during lexical ambiguity repetition in Parkinson's disease. *Cortex*, 45(8), 943-949.

Costafreda, S. G., Fu, C. H., Lee, L., Everitt, B., Brammer, M. J., & David, A. S. (2006). A Systematic Review and Quantitative Appraisal of fMRI Studies of Verbal Fluency: Role of the Left Inferior Frontal Gyrus. *Human Brain Mapping*, 27, 799-810.

Damen, E. J. P., & Brunia, C. H. M. (1987). Changes in Heart Rate and Slow Brain Potentials Related to Motor Preparation and Stimulus Anticipation in a Time Estimation Task. *Psychophysiology*, 24(6), 700–713. https://doi.org/10.1111/j.1469-8986.1987.tb00353.x

Damen, E. J. P., & Brunia, C. H. M. (1994). Is a stimulus conveying task-relevant information a sufficient condition to elicit a stimulus-preceding negativity? *Psychophysiology*, 31(2), 129–139. https://doi.org/10.1111/j.1469-8986.1994.tb01033.x

Damen, E. J. P., & Brunia, C. H. M. (1985). Slow brain potentials related to movement and visual feedback in a response timing task. *Biological Psychology*, 20(195).

DeLong, K. A., Chan, W., & Kutas, M. (2020). Testing limits: ERP evidence for word form preactivation during speeded sentence reading. *Psychophysiology*. https://doi.org/10.1111/psyp.13720

Delong, K. A., Troyer, M., & Kutas, M. (2014). Pre-Processing in Sentence Comprehension: Sensitivity to Likely Upcoming Meaning and Structure. *Linguistics and Language Compass*, 8(12), 631–645. https://doi.org/10.1111/lnc3.12093

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

Delong, K. A., Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology*, 48(9), 1203–1207. https://doi.org/10.1111/j.1469-8986.2011.01199.x

Delong, K.A., Quante, L., Kutas, M., (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia*, 61, 150–162.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9-21.

Dikker, S., & Pylkkänen, L. (2013). Predicting language: MEG evidence for lexical preactivation. *Brain and Language*, 127(1), 55–64. https://doi.org/10.1016/j.bandl.2012.08.004

Donchin, E., Ritter, W., & McCallum, C. (1978). Event-Related Brain Potentials in Man. https://doi.org/10.1016/B978-0-12-155150-6.X5001-9

Donkers, F. C. L., & van Boxtel, G. J. M. (2005). Mediofrontal negativities to averted gains and losses in the slot-machine task: A further investigation. *Journal of Psychophysiology*, 19(4), 256–262. https://doi.org/10.1027/0269-8803.19.4.256

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior research methods*, 45(4), 1246-1258. https://doi.org/10.3758/s13428-013-0326-1

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, 24(6), 409–436. https://doi.org/10.1007/BF02143160

Ehlrich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of verbal learning and verbal behavior*, 20(6), 641-655. doi: 10.1016/S0022-5371(81)90220-6

Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., & Mizuno, Y. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Movement Disorders*. 2007; 22: 1689–1707.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491-505. https://doi.org/10.1111/j.1469-8986.2007.00531.x

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, 41(4), 469-495. https://doi.org/10.1006/jmla.1999.2660

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, 42(2), 133–141. https://doi.org/10.1111/j.1469-8986.2005.00274.x

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and language*, 115(3), 149-161.

Federmeier, K. D., McLennan, D. B., Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, 39(2), 133–146. https://doi.org/10.1111/1469-8986.3920133

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007b). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146(1), 75–84. https://doi.org/10.1016/j.brainres.2006.06.101

Felser, C., Clahsen, H., & Münte, T. F. (2003). Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and wh-movement in German. *Brain and Language*, 87(3), 345-354. https://doi.org/10.1016/S0093-934X(03)00135-4

Fernández de Bobadilla, R., Pagonabarraga, J., Martínez-Horta, S., Pascual-Sedano, B., Campolongo, A., et al. (2013). Parkinson's disease-cognitive rating scale: Psychometrics for mild cognitive impairment. *Movement Disorders*, 28(10), 1376-1383.

Ferreira, F., & Chantavarin, S. (2018). Integration and Prediction in Language Processing: A Synthesis of Old and New. *Current Directions in Psychological Science*, 27(6), 443–448. https://doi.org/10.1177/0963721418794491

Ferreira, F., & Lowder, M. W. (2016). Prediction, information structure, and good-enough language processing. In Psychology of Learning and Motivation (Vol. 65, pp. 217-247). *Academic Press*. https://doi.org/10.1016/bs.plm.2016.04.002

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research*, 30(3), 321–338. https://doi.org/10.1023/A:1010447102554

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language*, 47(2), 250-272. https://doi.org/10.1016/S0749-596X(02)00004-9

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *The Quarterly Journal of Experimental Psychology Section*. 33(4), 465–495. https://doi.org/10.1080/14640748108400804

Forster, K. I. (1985). Lexical acquisition and the modular lexicon. *Language and Cognitive Processes*, 1(2), 87–108. https://doi.org/10.1080/01690968508402073

Foti, D., & Hajcak, G. (2012). Genetic variation in dopamine moderates neural response during reward anticipation and delivery: Evidence from event-related potentials. *Psychophysiology*, 49(5), 617–626. https://doi.org/10.1111/j.1469-8986.2011.01343.x

Foxe, J. J., Simpson, G. V., & Ahlfors, S. P. (1998). Parieto-occipital~ 1 0Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport*, 9(17), 3929-3933.

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, 140, 1–11. https://doi.org/10.1016/j.bandl.2014.10.006

Freunberger, D., & Roehm, D. (2016). Semantic prediction in language comprehension: evidence from brain potentials. *Language, Cognition and Neuroscience*, 31(9), 1193–1205. https://doi.org/10.1080/23273798.2016.1205202

Friederici, A. D., Kotz, S. A., Werheid, K., Hein, G., & von Cramon, D. Y. (2003). Syntactic processes in Parkinson's disease: a dissociation between early automatic and late integrational processes. *Neuropsychology*, 17(1), 133-142.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2010). The free-energy principle: a unified brain theory?, 11(2), 127–138. https://doi.org/10.1038/nrn2787

Fuentemilla, L., Cucurell, D., Marco-Pallarés, J., Guitart-Masip, M., Morís, J., & Rodríguez-Fornells, A. (2013). Electrophysiological correlates of anticipating improbable but desired events. *NeuroImage*, 78, 135–144. https://doi.org/10.1016/j.neuroimage.2013.03.062

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1), 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1

Goetz, C. G., Tilley, B.C., Shaftman, S.R., Stebbins, G.T., Fahn, S., Martinez-Martin, P., et al. (2008) Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders*; 23: 2129–2170.

Golob, E. J., Ovasapyan, V., & Starr, A. (2005). Event-related potentials accompanying motor preparation and stimulus expectancy in the young, young-old and oldest-old. *Neurobiology of aging*, 26(4), 531-542.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological review*, 101(3), 371.

Griffis, J. C., Nenert, R., Allendorfer, J. B., & Szaflarski, J. P. (2017). Damage to white matter bottlenecks contributes to language impairments after left hemispheric stroke. *NeuroImage: Clinical*, 14, 552-565.

Griffiths, B. J., Mayhew, S. D., Mullinger, K. J., Jorge, J., Charest, I., Wimber, M., & Hanslmayr, S. (2019). Alpha/beta power decreases track the fidelity of stimulus specific information. *ELife*, 8. https://doi.org/10.7554/eLife.49562

Grisoni, L., Miller, T. M., & Pulvermüller, F. (2017). Neural correlates of semantic prediction and resolution in sentence processing. *Journal of Neuroscience*, 2800-16. https://doi.org/10.1523/JNEUROSCI.2800-16.2017

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.

Grossman, M., A. C. Cooke, C. Devita, D. Alsop, J. Detre, W. Chen, et al. (2002). Age-related changes in working memory during sentence comprehension: an fMRI study. *Neuroimage*, 15.302–17.

Grossman, M., Cooke, A., Devita, C., Lee, C., Alsop, D., Detre, J., et al. (2003). Grammatical and resource components of sentence processing in Parkinson's disease: an fMRI study. *Neurology*, 60(5), 775-781.

Grünewald-Zuberbier, E., Grünewald, G., Runge, H., Netz, J., & Hömberg, V. (1981). Cerebral potentials during skilled slow positioning movements. *Biological Psychology*, 13(C), 71–87. https://doi.org/10.1016/0301-0511(81)90028-4

Hackley, S. A., Valle-Inclán, F., Masaki, H., & Hebert, K. (2014). Stimulus-preceding negativity (SPN) and attention to rewards. *Cognitive electrophysiology of attention: Signals of the mind*, 216-225. https://doi.org/10.1016/B978-0-12-398451-7.00017-8

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. *MIT Press*.

Halgren, E., Dhond, R. P., Christensen, N., Petten, C. Van, Marinkovic, K., Lewine, J. D., & Dale, A. M. (2002). N400-like Magnetoencephalography Responses Modulated by Semantic Context, Word Frequency, and Lexical Class in Sentences. *NeuroImage*, 17(3):1101-16. https://doi.org/10.1006/nimg.2002.1268

Händel, B. F., Haarmeier, T., & Jensen, O. (2011). Alpha oscillations correlate with the successful inhibition of unattended stimuli. *Journal of Cognitive Neuroscience*, 23(9), 2494–2502. https://doi.org/10.1162/jocn.2010.21557

Hanslmayr, S., Spitzer, B., & Bäuml, K. H. (2009). Brain oscillations dissociate between semantic and nonsemantic encoding of episodic memories. *Cerebral Cortex*, 19(7), 1631–1640. https://doi.org/10.1093/cercor/bhn197

Hanslmayr, S., Staudigl, T., & Fellner, M. C. (2012). Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Frontiers in human neuroscience*, 6, 74. https://doi.org/10.3389/fnhum.2012.00074

Hanslmayr, S., Volberg, G., Wimber, M., Raabe, M., Greenlee, M. W., & Bäuml, K. H. T. (2011). The relationship between brain oscillations and BOLD signal during memory formation: a combined EEG–fMRI study. *Journal of Neuroscience*, 31(44), 15674-15680.

Hart, J., & Gordon, B. (1990). Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Annals of Neurology*, 27(3), 226–231. https://doi.org/10.1002/ana.410270303

Heil, M., Rösler, F., & Hennighausen, E. (1996). Topographically distinct cortical activation in episodic long-terra memory: The retrieval of spatial versus verbal information. *Memory and Cognition*, 24(6), 777–795. https://doi.org/10.3758/BF03201102

Holcomb, P. J., & Grainger, J. (2006). On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience*, 18(10), 1631–1643. https://doi.org/10.1162/jocn.2006.18.10.1631

Holm, S. (1979) A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. 6, 65-70.

Horta-Barba, A., Pagonabarraga, J., Martínez-Horta, S., Marín-Lahoz, J., Sampedro, F., Fernández-Bobadilla, R., et al. (2020) The Free and Cued Selective Reminding Test in Parkinson's Disease Mild Cognitive Impairment: Discriminative Accuracy and Neural Correlates. *Frontiers in Neurology*. 11: 240.

Huang G., & Zhang, H. (2017) Improving sensitivity of cluster-based permutation test for EEG/MEG data, *8th International IEEE/EMBS Conference on Neural Engineering (NER)*, Shanghai, pp. 9-12, doi: 10.1109/NER.2017.8008279.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

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

Hughes A. J., Daniel S.E., Kilford L., &, Lees A.J. (1992) Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *Journal of Neurological Neurosurgery and Psychiatry*, 55:181-184.

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, 86, 157–171. https://doi.org/10.1016/j.jml.2015.10.007

Jackendoff, R. (2001). Foundations of Language, Brain, Meaning, Grammar, Evolution. *Behavioral and Brain Sciences*, 26, 651-707.

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

Jensen, O., Bonnefond, M., & Vanrullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends in Cognitive Sciences*, 16(4), 200–206. https://doi.org/10.1016/j.tics.2012.03.002

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, 4(1), 71–115. https://doi.org/10.1016/S0364-0213(81)80005-5

Kaan, E., & Carlisle, E. (2014). ERP Indices of Stimulus Prediction in Letter Sequences. *Brain sciences*, 4(4), 509-531. doi: 10.3390/brainsci4040509

Kabbara, A., Khalil, M., O'Neill, G., Dujardin, K., El Traboulsi, Y., Wendling, F., et al. (2019) Detecting modular brain states in rest and task. *Network Neuroscience*. Jul 1;3(3):878-901.

Kamide, Y. (2008). Anticipatory Processes in Sentence Processing. *Linguistics and Language Compass*, 2(4), 647–670. https://doi.org/10.1111/j.1749-818X.2008.00072.x

Kerlin, J. R., Shahin, A. J., & Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a "cocktail party." *Journal of Neuroscience*, 30(2), 620–628. https://doi.org/10.1523/JNEUROSCI.3631-09.2010

Khader, P. H., & Rösler, F. (2011). EEG power changes reflect distinct mechanisms during long-term memory retrieval. *Psychophysiology*, 48(3), 362–369. https://doi.org/10.1111/j.1469-8986.2010.01063.x

Khader, P. H., Jost, K., Ranganath, C., & Rösler, F. (2010). Theta and alpha oscillations during working-memory maintenance predict successful long-term memory encoding. *Neuroscience Letters*, 468(3), 339–343. https://doi.org/10.1016/j.neulet.2009.11.028

Khader, P., Ranganath, C., Seemüller, A., & Rösler, F. (2007). Working memory maintenance contributes to long-term memory formation: evidence from slow event-related brain potentials. *Cognitive, Affective & Behavioral Neuroscience*, 7(3), 212–224. https://doi.org/10.3758/CABN.7.3.212

Khader, P., Schicke, T., Röder, B., & Rösler, F. (2008). On the relationship between slow cortical potentials and BOLD signal changes in humans. *International Journal of Psychophysiology*, 67(3), 252–261. https://doi.org/10.1016/j.ijpsycho.2007.05.018

Kiehl, K. A., Laurens, K. R., & Liddle, P. F. (2002). Reading anomalous sentences: An event-related fMRI study of semantic processing. *NeuroImage*, 17(2), 842–850. https://doi.org/10.1016/S1053-8119(02)91244-9

Kielar, A., Panamsky, L., Links, K. A., & Meltzer, J. A. (2015). Localization of electrophysiological responses to semantic and syntactic anomalies in language comprehension with MEG. *Neuroimage*, 105, 507-524.

King, J. W., & Kutas, M. (1995). Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of cognitive neuroscience*, 7(3), 376-395. https://doi.org/10.1162/jocn.1995.7.3.376

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30(5), 580–602. https://doi.org/10.1016/0749-596X(91)90027-H

Kischka, U., Kammer, T. H., Maier, S., Weisbrod, M., Thimm, M., & Spitzer, M. (1996). Dopaminergic modulation of semantic network activation. *Neuropsychologia*, 34(11), 1107-1113.

Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology*, 24(1–2), 61–100. https://doi.org/10.1016/S0167-8760(96)00057-8

Klimesch, W. (1997). EEG-alpha rhythms and memory processes. *International Journal of psychophysiology*, 26(1-3), 319-340.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*. https://doi.org/10.1016/S0165-0173(98)00056-3

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2012.10.007

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

Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, 5(2), 196–214. https://doi.org/10.1162/jocn.1993.5.2.196

Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv Für Die Gesamte Physiologie Des Menschen Und Der Tiere*, 284(1), 1–17. https://doi.org/10.1007/BF00412364

Kotani, Y., & Aihara, Y. (1999). The effect of stimulus discriminability on stimulus-preceding negativities prior to instructive and feedback stimuli. *Biological Psychology*, 50(1), 1–18. https://doi.org/10.1016/S0301-0511(98)00047-7

Kotani, Y., Hiraku, S., Suda, K., & Aihara, Y. (2001). Effect of positive and negative emotion on stimulus-preceding negativity prior to feedback stimuli. *Psychophysiology*, 38(6), 873–878. https://doi.org/10.1111/1469-8986.3860873

Kotani, Y., Kishida, S., Hiraku, S., Suda, K., Ishii, M., & Aihara, Y. (2003). Effects of information and reward on stimulus-preceding negativity prior to feedback stimuli. *Psychophysiology*, 40(5), 818–826. https://doi.org/10.1111/1469-8986.00082

Kotani, Y., Ohgami, Y., Ishiwata, T., Arai, J. ichirou, Kiryu, S., & Inoue, Y. (2015). Source analysis of stimulus-preceding negativity constrained by functional magnetic resonance imaging. *Biological Psychology*, 111, 53–64. https://doi.org/10.1016/j.biopsycho.2015.08.005

Kotz, S. A., Frisch, S., Von Cramon, D. Y., & Friederici, A. D. (2003). Syntactic language processing: ERP lesion data on the role of the basal ganglia. *Journal of the International Neuropsychological Society*, 9(7), 1053-1060.

Kudlicka, A., Clare, L., & Hindle, J. V. (2011). Executive functions in Parkinson's disease: Systematic review and meta-analysis. *Movement disorders*, 26(13), 2305-2315.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

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.

Kuperberg, G.R. (2013). The proactive comprehender: What event-related potentials tell us about the dynamics of reading comprehension. In Miller, B., Cutting, L., & McCardle, P (Eds): Unraveling the Behavioral, Neurobiological, and Genetic Components of Reading Comprehension (pp. 176-19).

Kutas, M. (1997). Views on how the electrical activity that the brain generates reflects the functions of different language structures. *Psychophysiology*, 34(4), 383-398. https://doi.org/10.1111/j.1469-8986.1997.tb02382.x

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, 62(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205. Https://doi.org/10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. Nature, 307(5947), 161. Https://doi.org/10.1038/307161a0

Kutas, M., & King, J. W. (1996). 20 The Potentials for Basic Sentence Processing: Differentiating Integrative Processes.

Kutas, M., & King, J. W. (1996). The potentials for basic sentence processing: Differentiating integrative processes. *Attention and Performance*. XVI, 501–546.

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), Predictions in the brain: Using our past to generate a future (p. 190–207). *Oxford University Press*. https://doi.org/10.1093/acprof:oso/9780195395518.003.0065

Kutas, M., Iragui, V. J., Niu, Y. Q., D'Avanzo, T. J., Yang, J. C., Salmon, D. P., et al. (2013). Altered N400 congruity effects in Parkinson's disease without dementia. *Cognitive electrophysiology of attention: Signals of the mind*, 254-267.

Kutas, M., Van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Clinical Neurophysiology*, 69(3), 218-233. https://doi.org/10.1016/0013-4694(88)90131-9

Lakens, D. (2016). Will knowledge about more efficient study designs increase the willingness to pre-register? Retrieved from osf.io/uy8x5.

Lang, S., Hanganu A, Gan L., S., Kibreab, M., Auclair-Ouellet, N., Alrazi, T., et al. (2019) Network basis of the dysexecutive and posterior cortical cognitive profiles in Parkinson's disease. *Movement Disorders*. 34(6):893-902.

Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language*, 61(3), 326–338. https://doi.org/10.1016/j.jml.2009.06.004

Lau, E. F., & Namyst, A. (2019). fMRI evidence that left posterior temporal cortex contributes to N400 effects of predictability independent of congruity. *Brain and Language*, 199. https://doi.org/10.1016/j.bandl.2019.104697

Lau, E. F., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2013). Automatic semantic facilitation in anterior temporal cortex revealed through multimodal neuroimaging. *Journal of Neuroscience*, 33(43), 17174–17181. https://doi.org/10.1523/JNEUROSCI.1018-13.2013

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience*, 25(3), 484–502. https://doi.org/10.1162/jocn_a_00328

Lau, E. F., Namyst, A., Fogel, A., & Delgado, T. (2016). A Direct Comparison of N400 Effects of Predictability and Incongruity in Adjective-Noun Combination. *Collabra*, 2(1). https://doi.org/10.1525/collabra.40

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933. https://doi.org/10.1038/nrn2532

Lau, E. F., & Nguyen, E. (2015). The role of temporal predictability in semantic expectation: An MEG investigation. *Cortex*, 68, 8-19. https://doi.org/10.1016/j.cortex.2015.02.022

León-Cabrera, P., Flores, A., Rodríguez-Fornells, A., & Morís, J. (2019). Ahead of time: Early sentence slow cortical modulations associated to semantic prediction. *NeuroImage*, 189, 192-201.

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

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

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

Li, X., Zhang, Y., Xia, J., & Swaab, T. Y. (2017). Internal mechanisms underlying anticipatory language processing: Evidence from event-related-potentials and neural oscillations. *Neuropsychologia*, 102, 70-81.

Li, X., Zhang, Y., Xia, J., & Swaab, T. Y. (2017). Neuropsychologia Internal mechanisms underlying anticipatory language processing: Evidence from event-related-potentials and neural oscillations. *Neuropsychologia*, 102(16), 70–81. https://doi.org/10.1016/j.neuropsychologia.2017.05.017

Litvan, I., Aarsland, D., Adler, C. H., Goldman, J. G., Kulisevsky, J., Mollenhauer, B., et al. (2011). MDS Task Force on mild cognitive impairment in Parkinson's disease: critical review of PD-MCI. *Movement disorders*, 26(10), 1814-1824

López-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*, 8, 213. Https://doi.org/10.3389/fnhum.2014.00213

Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology*, 88, 22-60. https://doi.org/10.1016/j.cogpsych.2016.06.002

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

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8(1), 1-71. doi: 10.1016/0010-0277(80)90015-3

Martinez-Horta, & S., Kulisevsky, J. (2019). Mild cognitive impairment in Parkinson's disease. *Journal of Neural Transmission*. 126(7), 897-904.

Mathewson, K. E., Beck, D. M., Ro, T., Maclin, E. L., Low, K. A., Fabiani, M., & Gratton, G. (2014). Dynamics of alpha control: Preparatory suppression of posterior alpha oscillations by frontal modulators revealed with combined EEG and event-related optical signal (EROS). *Journal of Cognitive Neuroscience*, 26(10), 2400–2415. https://doi.org/10.1162/jocn_a_00637

Mattox, S. T., Valle-Inclán, F., & Hackley, S. A. (2006). Psychophysiological evidence for impaired reward anticipation in Parkinson's disease. *Clinical Neurophysiology*, 117(10), 2144-2153.

Matzke, M., Mai, H., Nager, W., Rüsseler, J., & Münte, T. (2002). The costs of freedom: An ERP - Study of non-canonical sentences. *Clinical Neurophysiology*, 113(6), 844–852. https://doi.org/10.1016/S1388-2457(02)00059-7

Mazaheri, A., van Schouwenburg, M. R., Dimitrijevic, A., Denys, D., Cools, R., & Jensen, O. (2014). Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities. *NeuroImage*, 87, 356–362. https://doi.org/10.1016/j.neuroimage.2013.10.052

McDonald, S. A., & Shillcock, R. C. (2003). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, 43, 1735–1751. doi: 10.1016/S0042-6989(03)00237-2

Meeuwissen, E. B., Takashima, A., Fernández, G., & Jensen, O. (2011). Increase in posterior alpha activity during rehearsal predicts successful long-term memory formation of word sequences. *Human Brain Mapping*, 32(12), 2045-2053.

Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2006). Watching the brain during meaning acquisition. *Cerebral Cortex*, 17(8), 1858-1866. https://doi.org/10.1093/cercor/bhl094

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *European Journal of Neuroscience*, 48(7), 2609–2621. https://doi.org/10.1111/ejn.13748

Michalewski, H. J., Thompson, L. W., Smith, D. B., Patterson, J. V., Bowman, T. E., Litzelman, D., et al. (1980). Age differences in the contingent negative variation (CNV): reduced frontal activity in the elderly. *Journal of Gerontology*, 35(4), 542-549.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*. https://doi.org/10.1146/annurev.neuro.24.1.167

Molinaro, N., Monsalve, I. F., & Lizarazu, M. (2016). Is there a common oscillatory brain mechanism for producing and predicting language?, *Language, Cognition and Neuroscience*, 31:1, 145-158, https://doi.org/10.1080/23273798.2015.1077978

Monchi, O., Petrides, M., Doyon, J., Postuma, R.B., Worsley, K., Dagher, A., 2004. Neural bases of set-shifting deficits in Parkinson's disease. *Journal of Neuroscience*. 24, 702–710. http://dx.doi.org/10.1523/JNEUROSCI.4860-03.2004.

Morís, J., Luque, D., & Rodríguez-Fornells, A. (2013). Learning-induced modulations of the stimulus-preceding negativity. *Psychophysiology*, 50(9), 931-939. https://doi.org/10.1111/psyp.12073

Morris, J., Grainger, J., & Holcomb, P. (2008). An electrophysiological investigation of early effects of masked morphological priming. *Language and Cognitive Processes*, 23(7–8), 1021–1056. https://doi.org/10.1080/01690960802299386

Morris, R. K. (2006). Lexical processing and sentence context effects. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics*. 2nd ed., pp. 377 – 402. doi: 10.1016/b978-012369374-7/50011-0

Müller, H. M., King, J. W., & Kutas, M. (1997a). Event-related potentials elicited by spoken relative clauses. *Cognitive Brain Research*, 5(3), 193–203. https://doi.org/10.1016/S0926-6410(96)00070-5

Münte, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature*, 395(6697), 71–73. https://doi.org/10.1038/25731

Muslimović, D., Post, B., Speelman, J. D., & Schmand, B. (2005). Cognitive profile of patients with newly diagnosed Parkinson disease. *Neurology*, 65(8), 1239-1245.

Newman, R. L., & Connolly, J. F. (2009). Electrophysiological markers of pre-lexical speech processing: Evidence for bottom-up and top-down effects on spoken word processing. *Biological Psychology*, 80(1), 114–121. https://doi.org/10.1016/j.biopsycho.2008.04.008

Nieuwland, M. S. (2015). The truth before and after: Brain potentials reveal automatic activation of event knowledge during sentence comprehension. *Journal of Cognitive Neuroscience*, 27(11), 2215-2228. doi: 10.1162/jocn_a_00856

Nieuwland, M. S., & Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of cognitive neuroscience*, 18(7), 1098-1111. https://doi.org/10.1162/jocn.2006.18.7.1098

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E, Donaldson, D. I., et al. (2020). Dissociable effects of prediction and integration during language comprehension: evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society B*. B37520180522. http://doi.org/10.1098/rstb.2018.0522

Ohgami, Y., Kotani, Y., Hiraku, S., Aihara, Y., & Ishii, M. (2004). Effects of reward and stimulus modality on stimulus-preceding negativity. *Psychophysiology*, 41, 729–738. doi: 10.1111/j.1469-8986.2004.00203.x

Olichney, J. M., Taylor, J. R., Gatherwright, J., Salmon, D. P., Bressler, A. J., Kutas, M., et al. (2008). Patients with MCI and N400 or P600 abnormalities are at very high risk for conversion to dementia. *Neurology*, 70(19 Part 2), 1763-1770.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 1. https://doi.org/10.1155/2011/156869

Otten, M., & Van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, 45(6), 464–496. https://doi.org/10.1080/01638530802356463

Pagonabarraga, J., & Kulisevsky, J. (2012). Cognitive impairment and dementia in Parkinson’s disease. *Neurobiology of Disease*.

Pagonabarraga, J., Kulisevsky, J., Llebaria, G., García-Sánchez, C., Pascual-Sedano, B., & Gironell, A. (2008). Parkinson's Disease-Cognitive Rating Scale: A New Cognitive Scale Specific for Parkinson's disease. *Movement disorders: Official Journal of the Movement Disorder Society*, 23(7), 998-1005.

Paxton, J. L., Barch, D. M., Racine, C. A., & Braver, T. S. (2008). Cognitive control, goal maintenance, and prefrontal function in healthy aging. *Cerebral Cortex* (New York, N.Y.: 1991), 18(5), 1010–1028. https://doi.org/10.1093/cercor/bhm135

Payne, B. R., & Federmeier, K. D. (2018). Contextual constraints on lexico-semantic processing in aging: Evidence from single-word event-related brain potentials. *Brain Research*, 1687, 117-128.

Payne, B. R., & Silcox, J. W. (2019). Aging, context processing, and comprehension. In Psychology of learning and motivation (Vol. 71, pp. 215-264). *Academic Press*.

Pell, M., & Monetta, L. (2008). How Parkinson's disease affects non-verbal communication and language processing. *Language and Linguistics Compass*, 2 (5): 739–759.

Pezzulo, G., Butz, M. V., & Castelfranchi, C. (2008). The anticipatory approach: definitions and taxonomies. In The challenge of anticipation. *Springer*, Berlin, Heidelberg (pp. 23-43).

Piai, V., & Zheng, X. (2019). Speaking waves: Neuronal oscillations in language production. In Psychology of Learning and Motivation - Advances in Research and Theory (Vol. 71, pp. 265–302). *Academic Press Inc*. https://doi.org/10.1016/bs.plm.2019.07.002

Piai, V., Dahlslätt, K., & Maris, E. (2015). Statistically comparing EEG/MEG waveforms through successive significant univariate tests: How bad can it be?. *Psychophysiology*, 52(3), 440-443. https://doi.org/10.1111/psyp.12335

Piai, V., Klaus, J., & Rossetto, E. (2020). The lexical nature of alpha-beta oscillations in context-driven word production. *Journal of Neurolinguistics*, 55, 100905.

Piai, V., Meyer, L., Dronkers, N. F., & Knight, R. T. (2017). Neuroplasticity of language in left-hemisphere stroke: Evidence linking subsecond electrophysiology and structural connections. *Human Brain Mapping*, 38(6), 3151–3162. https://doi.org/10.1002/hbm.23581

Piai, V., Meyer, L., Schreuder, R., & Bastiaansen, M. C. (2013). Sit down and read on: Working memory and long-term memory in particle-verb processing. *Brain and language*, 127(2), 296-306. https://doi.org/10.1016/j.bandl.2013.09.015

Piai, V., Roelofs, A., & Maris, E. (2014). Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. *Neuropsychologia*, 53, 146-156. https://doi.org/10.1016/j.neuropsychologia.2013.11.014

Piai, V., Roelofs, A., Rommers, J., & Maris, E. (2015). Beta oscillations reflect memory and motor aspects of spoken word production. *Human Brain Mapping*, 36(7), 2767–2780. https://doi.org/10.1002/hbm.22806

Piai, V., Rommers, J., & Knight, R. T. (2018). Lesion evidence for a critical role of left posterior but not frontal areas in alpha-beta power decreases during context-driven word production. *European Journal of Neuroscience*, 48(7), 2622–2629. https://doi.org/10.1111/ejn.13695

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. https://doi.org/10.1037/bul0000158

Poli, S., Sarlo, M., Bortoletto, M., Buodo, G., & Palomba, D. (2007). Stimulus-Preceding Negativity and heart rate changes in anticipation of affective pictures. *International Journal of Psychophysiology*, 65(1), 32–39. https://doi.org/10.1016/j.ijpsycho.2007.02.008

Politzer-Ahles, S., Xiang, M., & Almeida, D. (2017). “Before” and “after”: Investigating the relationship between temporal connectives and chronological ordering using eventrelated potentials. *PLoS ONE*, 12(4), 1–17. https://doi.org/10.1371/journal.pone.0175199

Posner, M. I., & Petersen, S. E. (1990). The Attention System of the Human Brain. *Annual Review of Neuroscience*, 13(1), 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Prystauka, Y., & Lewis, A. G. (2019). The power of neural oscillations to inform sentence comprehension: A linguistic perspective. *Language and Linguistics Compass*, 13(9), e12347.

Pulvermüller, F., & Grisoni, L. (2020). Semantic Prediction in Brain and Mind. *Trends in Cognitive Sciences*, 24(10), 781-784. https://doi.org/10.1016/j.tics.2020.07.002

Pylkkänen, L., Oliveri, B., & Smart, A. J. (2009). Semantics vs. world knowledge in prefrontal cortex. *Language and Cognitive Processes*, 24(9), 1313–1334. https://doi.org/10.1080/01690960903120176

Rabovsky, M. (2020). Change in a probabilistic representation of meaning can account for N400 effects on articles: A neural network model. *Neuropsychologia*, 143, 107466. https://doi.org/10.1016/j.neuropsychologia.2020.107466

Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*, 2(9), 693–705. https://doi.org/10.1038/s41562-018-0406-4

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye Movements and Word Skipping During Reading: Effects of Word Length and Predictability. *Journal of Experimental Psychology: Human Perception and Performance*, 37(2), 514–528. https://doi.org/10.1037/a0020990

Ren, X., Valle-Inclán, F., Tukaiev, S., & Hackley, S. A. (2017). Changes in the stimulus-preceding negativity and lateralized readiness potential during reinforcement learning. *Psychophysiology*, 54(7), 969–981. https://doi.org/10.1111/psyp.12859

Richard Jennings, J., & Wood, C. C. (1976). The ε-adjustment procedure for repeated-measures analyses of variance. *Psychophysiology*, 13(3), 277-278. https://doi.org/10.1111/j.1469-8986.1976.tb00116.x

Rihs, T. A., Michel, C. M., & Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by α-band EEG synchronization. *European Journal of Neuroscience*, 25(2), 603–610. https://doi.org/10.1111/j.1460-9568.2007.05278.x

Rodríguez-Fornells, A., Kurzbuch, A. R., & Münte, T. F. (2002). Time course of error detection and correction in humans: neurophysiological evidence. *The Journal of Neuroscience*, 22(22), 9990-9996.

Rohrbaugh, J. W., & Gaillard, A. W. K. (1983). 13 Sensory and Motor Aspects of the Contingent Negative Variation. *Advances in Psychology*, 10(C), 269–310. https://doi.org/10.1016/S0166-4115(08)62044-0

Rolke, B., Heil, M., Hennighausen, E., Häussler, C., & Rösler, F. (2000). Topography of brain electrical activity dissociates the sequential order transformation of verbal versus spatial information in humans. *Neuroscience Letters*, 282(1–2), 81–84. https://doi.org/10.1016/S0304-3940(00)00891-0

Rommers, J., Dickson, D. S., Norton, J. J., 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

Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the Human brain: Evidence from idiom comprehension. *Journal of Cognitive Neuroscience*, 25(5), 762–776. https://doi.org/10.1162/jocn_a_00337

Roos, N. M., & Piai, V. (2020). Across-session consistency of context-driven language processing: A magnetoencephalography study. *The European journal of neuroscience*, 52(5), 3457–3469. https://doi.org/10.1111/ejn.14785

Rösler, F., Heil, M., & Hennighausen, E. (1995). Distinct cortical activation patterns during long-term memory retrieval of verbal, spatial, and color information. *Journal of Cognitive Neuroscience*, 7(1), 51–65. https://doi.org/10.1162/jocn.1995.7.1.51

Rösler, F., Heil, M., & Hennighausen, E. (1995). Exploring memory functions by means of brain electrical topography: a review. *Brain Topography*, 7(4), 301-313. https://doi.org/10.1007/BF01195256

Rösler, F., Heil, M., & Röder, B. (1997). Slow negative brain potentials as reflections of specific modular resources of cognition. In Biological Psychology (Vol. 45, pp. 109–141). https://doi.org/10.1016/S0301-0511(96)05225-8

Ruchkin, D. S., Johnson Jr, R., Canoune, H., & Ritter, W. (1990). Short-term memory storage and retention: An event-related brain potential study. *Electroencephalography and clinical Neurophysiology*, 76(5), 419-439. https://doi.org/10.1016/0013-4694(90)90096-3

Ruchkin, D. S., Sutton, S., Mahaffey, D., & Glaser, J. (1986). Terminal CNV in the absence of motor response. *Electroencephalography and Clinical Neurophysiology*, 63(5), 445–463. https://doi.org/10.1016/0013-4694(86)90127-6

Ruchkin, Daniel S., Berndt, R. S., Johnson, R., Ritter, W., Grafman, J., & Canoune, H. L. (1997). Modality-specific processing streams in verbal working memory: Evidence from spatio-temporal patterns of brain activity. *Cognitive Brain Research*, 6(2), 95–113. https://doi.org/10.1016/S0926-6410(97)00021-9

Ruchkin, Daniel S., Johnson, R., Canoune, H., & Ritter, W. (1990). Short-term memory storage and retention: an event-related brain potential study. *Electroencephalography and Clinical Neurophysiology*, 76(5), 419–439. https://doi.org/10.1016/0013-4694(90)90096-3

Sampedro, F., Marín-Lahoz, J., Martínez-Horta, S., Pagonabarraga, J., & Kulisevsky, J. (2019). Dopaminergic degeneration induces early posterior cortical thinning in Parkinson's disease. *Neurobiology of Disease*, 124, 29–35. https://doi.org/10.1016/j.nbd.2018.11.001

Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic Relatedness and the Scope of Facilitation for Upcoming Words in Sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(2), 344–354. https://doi.org/10.1037/0278-7393.14.2.344

Schwanenflugel, P. J., & Shoben, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language*, 24(2), 232–252. https://doi.org/10.1016/0749-596X(85)90026-9

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition* (Vol. 71).

Service, E., Helenius, P., Maury, S., & Salmelin, R. (2007). Localization of syntactic and semantic brain responses using magnetoencephalography. *Journal of Cognitive Neuroscience*, 19(7), 1193–1205. https://doi.org/10.1162/jocn.2007.19.7.1193

Singer, M., Graesser, A. C., & Trabasso, T. (1994). Minimal or global inference during reading. *Journal of Memory and Language*, 33(4), 421-441.

Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, 128(3), 302-319. doi: 10.1016/j.cognition.2013.02.013

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology* (Vol. 45).

Staub, A. (2015). The Effect of Lexical Predictability on Eye Movements in Reading: Critical Review and Theoretical Interpretation. *Language and Linguistics Compass*, 9(8), 311–327. https://doi.org/10.1111/lnc3.12151

Stavropoulos, K. K., & Carver, L. J. (2014). Reward anticipation and processing of social versus nonsocial stimuli in children with and without autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 55(12), 1398-1408. doi: 10.1111/jcpp.12270

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences,* 3(4), 151-162.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience*, 16(13), 4240–4249. https://doi.org/10.1523/jneurosci.16-13-04240.1996

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Taylor, W. L. (1953). "Cloze Procedure": A New Tool for Measuring Readability. *Journalism Quarterly*, 30(4), 415–433. https://doi.org/10.1177/107769905303000401

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.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America*, 94(26), 14792–14797. https://doi.org/10.1073/pnas.94.26.14792

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

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). α-Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *The Journal of Neuroscience*, 26(37), 9494–9502. doi:10.1523/JNEUROSCI.0875-06.2006

Tomlinson, C. L., Stowe, R., Patel, S., Rick, C., Gray, R., Clarke, C. E. (2010) Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Movement Disorders*; 25: 2649–2653.

Traxler, M. J., & Foss, D. J. (2000). Effects of Sentence Constraint on Priming in Natural Language Comprehension. *Journal of Experimental Psychology: Learning Memory and Cognition*, 26(5), 1266–1282. https://doi.org/10.1037/0278-7393.26.5.1266

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *Quarterly Journal of Experimental Psychology*, 64(3), 447-466.

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning Memory and Cognition*, 31(3), 443–467. https://doi.org/10.1037/0278-7393.31.3.443

van Berkum, J. J., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, 41(2), 147-182. https://doi.org/10.1006/jmla.1999.2641

van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443. https://doi.org/10.1037/0278-7393.31.3.443

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

van Boxtel, G. J., & Böcker, K. B. E. (2004). Cortical measures of anticipation. *Journal of Psychophysiology*, 18, 61–76. doi: 10.1027/0269-8803.18.23.61

van Boxtel, G. J.M., & Brunia, C. H. M. (1994). Motor and non-motor aspects of slow brain potentials. *Biological Psychology*, 38(1), 37–51. https://doi.org/10.1016/0301-0511(94)90048-5

van der Molen, M. J., Poppelaars, E. S., Van Hartingsveldt, C. T., Harrewijn, A., Gunther Moor, B., & Westenberg, P. M. (2013). Fear of negative evaluationmodulates electrocortical and behavioral responses when anticipating socialevaluative feedback. *Frontiers in Human Neuroscience*, 7, 936. doi: 10.3389/fnhum.2013.00936

van Dijk, H., Schoffelen, J.M., Oostenveld, R., & Jensen, O. (2008). Pre-stimulus oscillatory activity in the alpha band predicts visual discrimination ability. Journal of Neuroscience. 28, 1816–1823.doi:10.1523/jneurosci.1853- 07.2008

van Dijk, T. A., & Kintsch, W. (1982). Strategies of Discourse Comprehension.

van Ede, F. (2018). Mnemonic and attentional roles for states of attenuated alpha oscillations in perceptual working memory: a review. *European Journal of Neuroscience*, 48(7), 2509–2515. https://doi.org/10.1111/EJN.13759@10.1111/(ISSN)1460-9568.BRITISH_NEUROSCIENCE_ASSOCIATION_FESTIVAL_OF_SCIENCE2019

van Ede, F. (2018). Mnemonic and attentional roles for states of attenuated alpha oscillations in perceptual working memory: a review. *European Journal of Neuroscience*, 48(7), 2509-2515.

van Ede, F., Niklaus, M., & Nobre, A. C. (2017). Temporal expectations guide dynamic prioritization in visual working memory through attenuated α oscillations. *Journal of Neuroscience*, 37(2), 437–445. https://doi.org/10.1523/JNEUROSCI.2272-16.2016

van Petten, C. V., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience*, 3(2), 131-150. doi: 10.1162/jocn.1991.3.2.131

van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open-and closed-class words. *Memory & Cognition*, 19(1), 95-112. https://doi.org/10.3758/BF03198500

van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, 97(3), 279–293. https://doi.org/10.1016/j.bandl.2005.11.003

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 Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia*, 33(4), 485–508. https://doi.org/10.1016/0028-3932(94)00133-A

van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(2), 394. doi: 10.1037/0278-7393.25.2.394

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

Vos, S. H., Gunter, T. C., Schriefers, H., & Friederici, A. D. (2001). Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. *Language and Cognitive Processes*, 16(1), 65-103. https://doi.org/10.1080/01690960042000085

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, 31(2), 329–338. https://doi.org/10.1016/S0896-6273(01)00359-2

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensori-motor association and expectancy in the human brain. *Nature*, 203(4943), 380–384. https://doi.org/10.1038/203380a0

Wang, L., Hagoort, P., & Jensen, O. (2018). Gamma Oscillatory Activity Related to Language Prediction. *Journal of Cognitive Neuroscience*, 1-12. https://doi.org/10.1162/jocn_a_01275

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.

Wang, L., Kuperberg, G., & Jensen, O. (2018). Specific lexico-semantic predictions are associated with unique spatial and temporal patterns of neural activity. *ELife*, 7. https://doi.org/10.7554/eLife.39061

Wang, L., Zhu, Z., & Bastiaansen, M. C. M. (2012). Integration or predictability? A further specification of the functional role of gamma oscillations in language comprehension. *Frontiers in Psychology*, 3, 187. doi:10.3389/fpsyg.2012.00187

Weisz, N., Hartmann, T., Müller, N., & Obleser, J. (2011). Alpha rhythms in audition: cognitive and clinical perspectives. *Frontiers in Psychology*, 2, 73. https://doi.org/10.3389/fpsyg.2011.00073

Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., et al. (2016). Verbal fluency: Language or executive function measure?. *Applied Neuropsychology: Adult*, 23(1), 29-34. doi: 10.1080/23279095.2015.1004574.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, 16(7), 1272–1288. https://doi.org/10.1162/0898929041920487

Wicha, N. Y., Bates, E. A., Moreno, E. M., & Kutas, M. (2003). Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience letters*, 346(3), 165-168. https://doi.org/10.1016/S0304-3940(03)00599-8

Willems, R. M., Frank, S. L., Nijhof, A. D., Hagoort, P., & Bosch, A. Van Den. (2016). Prediction During Natural Language Comprehension. *Cerebral Cortex*, 26 (6), 2506–2516, https://doi.org/10.1093/cercor/bhv075

Wlotko, E. W., & Federmeier, K. D. (2012). Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension. *Psychophysiology*, 49(6), 770-785.

Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage*, 62(1), 356–366. https://doi.org/10.1016/j.neuroimage.2012.04.054

Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: Age-related differences in the use of sentential context. *Psychology and Aging*, 27(4), 975–988. https://doi.org/10.1037/a0029206

Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific α-bank electroencephalography increases over occipital cortex. *Journal of Neuroscience*, 20(6), RC63-RC63. https://doi.org/10.1523/jneurosci.20-06-j0002.2000

Zacks, J. M., Mar, R. A., & Calarco, N. (2018). The cognitive neuroscience of discourse: Covered ground and new directions. In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), Routledge handbooks in linguistics. The Routledge handbook of discourse processes (p. 269–294).