Neuroscience and Biobehavioral Reviews 160 (2024) 105624

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Neurophysiological signatures of prediction in language: A critical review of anticipatory negativities

Patricia León-Cabrera [a,f], Anna Hjortdal [d], Sabine Gosselke Berthelsen [d,e], Antoni Rodríguez-Fornells [a,b,c], Mikael Roll [d,*]

[a] Cognition and Brain Plasticity Unit, Department of Cognition, Development and Educational Psychology, University of Barcelona, Barcelona, Spain
[b] Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Hospitalet de Llobregat, Barcelona, Spain
[c] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain
[d] Centre for Languages and Literature (SOL), Lund University, Lund, Sweden
[e] Department of Nordic Studies and Linguistics (NorS), University of Copenhagen, Copenhagen, Denmark
[f] Department of Basic Sciences, Universitat Internacional de Catalunya, Sant Cugat del Vallès, Barcelona, Spain

ARTICLE INFO

Keywords:
Event-related potentials
Prediction
Anticipatory negativity
Comprehension
Prediction negativity
Pre-activation negativity

ABSTRACT

Recent event-related potential (ERP) studies in language comprehension converge in finding anticipatory negativities preceding words or word segments that can be pre-activated based on either sentence contexts or phonological cues. We review these findings from different paradigms in the light of evidence from other cognitive domains in which slow negative potentials have long been associated with anticipatory processes and discuss their potential underlying mechanisms. We propose that this family of anticipatory negativities captures common mechanisms associated with the pre-activation of linguistic information both within words and within sentences. Future studies could utilize these anticipatory negativities in combination with other, well-established ERPs, to simultaneously track prediction-related processes emerging at different time intervals (before and after the perception of pre-activated input) and with distinct time courses (shorter-lived and longer-lived cognitive operations).

1. Introduction

A common research enterprise in cognitive neuroscience is understanding to what extent prediction occurs in the brain and how it is implemented at the neural level (Clark, 2013; Friston, 2010). A crucial part of this challenge is deciphering how our brains build up information ex-ante, prior to the advent of an incoming event. Although a large body of research exists in perceptual and motor domains, much remains to be uncovered in other fields, such as language. Prediction might be particularly relevant in human communication mediated by language, as it allows a more efficient, fluent, and fast intercommunication between speakers (Kutas et al., 2011; Pickering and Garrod, 2013). However, due to the high ambiguity and noise in the to-be-decoded signal, some researchers have pointed out that, to a certain extent, prediction in language could be rather limited or even unnecessary (e.g., Huettig and Mani, 2016). Therefore, ascertaining the role of prediction in language continues to be a meaningful research endeavor.

Recent research using *event-related brain potentials* (ERPs) has provided new evidence of elicited electrical brain responses associated with the prediction of linguistic information in comprehension (Grisoni et al., 2017; León-Cabrera et al., 2017, 2019, 2021; Roll et al., 2013, 2015, 2017; Söderström et al., 2016, 2018). Capitalizing on the fine-grained temporal resolution of cortical electrical activity, researchers have turned their attention to brain changes elicited *before* linguistic items (henceforth *targets*) that can be predicted from information in the context. This is an excellent addition to the otherwise current dominant approach in ERP research on linguistic prediction, which is mainly focused on investigating brain activity changes in the post-target interval, especially by means of amplitude changes in the N400 component (as reviewed in Lau et al., 2008). This post-target approach can tap into the processing of input that either matches or mismatches the predictions. However, it offers an incomplete picture of the full chain of processes involved in the endeavor of pre-activating linguistic representations. By definition, predictive mechanisms should already be at play in the pre-target interval —*before* the expected word (or other linguistic information) is perceived.

* Corresponding author.
E-mail addresses: leoncabrera.patricia@gmail.com (P. León-Cabrera), mikael.roll@ling.lu.se (M. Roll).

https://doi.org/10.1016/j.neubiorev.2024.105624
Received 22 June 2023; Received in revised form 9 February 2024; Accepted 13 March 2024
Available online 16 March 2024

This shift from post- to pre-target analysis has revealed that ERP signatures seem to converge on sustained negative potentials when the pre-activation of upcoming words is possible. The present review aims to comprehensively integrate these anticipatory sustained negativities with existing work and highlight their value in understanding predictive language processing. To this end, we first provide a brief overview of the study of prediction in the field of language comprehension, with an emphasis on post-target ERP components capturing the effects of successful and failed predictions. Second, we introduce important evidence from other cognitive domains (beyond language) that have consistently associated the appearance of slow negative potentials with the anticipation of upcoming, relevant target-stimuli. Then, we review the evidence of sustained negativities in language preceding targets which can be predicted from complex cues like sentence contexts or from simpler cues like phonemes. Finally, we discuss the functional significance of these anticipatory sustained negativities and their relation with other ERPs associated with prediction and anticipatory processing.

2. Prediction in language comprehension

Language comprehension is a complex endeavor that involves accessing and combining different types of information, ranging from specific orthographic or phonological representations to syntactic, semantic, and pragmatic information. In early models of language processing, the key mechanism supporting comprehension was thought to be integration, which assumed that meaning is incrementally constructed and updated as the linguistic input flows in (Marslen-Wilson, 1987; Norris, 1986). Building upon this predominantly bottom-up driven processing, interest has moved to top-down models of language processing that are highly dynamic (using all available information) (Hagoort, 2014) and that assume that meaning can get ahead of the input (Collins and Loftus, 1975; Lupyan and Clark, 2015; McClelland, 2013; McClelland and Elman, 1986). Nowadays, psycholinguistic theories underscore the importance of prediction for language processing (Dell and Chang, 2014; Kuperberg and Jaeger, 2016), which aligns with the broader conceptualization in the cognitive science of the brain as a "prediction machine" (Clark, 2013; Friston, 2010).

In language comprehension, prediction can be defined as the pre-activation of linguistic representations, before they have been partially or fully activated by the bottom-up input (Huettig et al., 2022; Kutas and Federmeier, 2011). Unlike activation triggered by bottom-up input, pre-activation is driven by the information in the prior context (e.g., words), which can pre-activate other, related representations through passive spreading activation in memory networks (Collins and Loftus, 1975; McClelland and Elman, 1986; Morton, 1969). In addition, higher level representations built upon the prior context may be used to more actively pre-activate specific representations (for discussions about different pre-activation mechanisms, see Huettig et al., 2022; Kuperberg and Jaeger, 2016). There is abundant empirical evidence showing that comprehenders can use diverse information in the prior context to pre-activate linguistic content at different representational levels. Ranging from broad predictions about upcoming events (discussed in Kuperberg, 2013; Kuperberg, 2021) down to the pre-activation of specific features, including semantic (Federmeier and Kutas, 1999; Wang, Kuperberg and Jensen, 2018) and morphosyntactic features (van Berkum et al., 2005; Wicha et al., 2004). Word-form information (phonological or orthographic) can also be pre-activated (DeLong et al., 2005, 2021; Laszlo and Federmeier, 2009, Salverda et al., 2014), although some studies indicate that this occurs only under certain circumstances, such as when there is sufficient time to do so (Freunberger and Roehm, 2016; Ito et al., 2016). The influence of factors like time (Huettig and Guerra, 2019) or task-related goals (Brothers et al., 2020) on word-form pre-activation has led some authors to conclude that prediction might not be a default operation in language comprehension (Huettig and Mani, 2016). For example, it has been proposed that word-form pre-activation is contingent on the formation of a sufficiently rich representation of meaning (Ferreira and Chantavarin, 2018). Moreover, several studies seem to suggest that some forms of linguistic prediction depend on having enough cognitive resources available (Federmeier, 2007; Federmeier et al., 2002; Huettig, 2015). Yet, although possible limits on linguistic prediction should be acknowledged, there now seems to exist a broad consensus that prediction is at work during language comprehension, at least under some conditions (Huettig et al., 2022; Kuperberg and Jaeger, 2016; Kutas and Federmeier, 2011).

Much of the aforementioned evidence regarding prediction in language comprehension has been obtained from ERPs, as their high temporal resolution is very convenient for examining fast-occurring and short-lasting language-related processes. A common strategy for studying prediction mechanisms in language involves measuring brain responses associated with linguistic items (here referred to as "targets") that either fulfill or violate contextual expectations. Temporarily, two intervals can be distinguished: the pre-target interval, which is more likely to capture processes associated with the generation of predictions, and the post-target interval, which should capture the benefits and costs of processing based on prediction. It should be clarified that a target can be an item (e.g., a lexical item, a phoneme) or part of an item (e.g., a part of a word, or a fragment of a sentence). Most of the work has focused on predictability effects on well-established language-related ERPs that arise in the post-target interval of words—specifically, the N400, the LAN, and the P600.

In psycholinguistic studies, the predictability of a word in its context is typically operationalized as cloze probability–the percentage of individuals that supply that word as a continuation of a particular sentence in an offline test (Taylor, 1953). For example, the sentence "Don't touch the wet" is continued with the word "paint" by most respondents, while only a few provide "dog" as a continuation. Thus, in this context, 'paint' has a higher cloze probability–i.e., is more predictable–than 'dog'. Using this measure, many studies on written language have shown that readers spend less time fixating words that are predictable and are more likely to skip them (Balota et al., 1985; Rayner et al., 2011; for a review, see Staub, 2015). Furthermore, behavioral 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 and LaCount, 1988; Schwanenflugel and Shoben, 1985; Traxler and Foss, 2000).

The N400 component is a centro-parietally distributed and negative-going component, which peaks approximately 400 ms after the onset of meaningful words, and is strongly linked to lexico-semantic processing (Kutas and Hillyard, 1980). Relevantly, the amplitude of the N400 component is inversely correlated with the cloze probability (i.e., predictability) of the eliciting word (Kutas and Hillyard, 1984): it is larger for unpredictable than for predictable words, suggesting facilitated lexical-semantic processing of the latter (see also, Federmeier, 2007; Federmeier and Kutas, 1999; Federmeier et al., 2007; van Petten et al., 1999; Wlotko and Federmeier, 2012). A long-standing debate is whether the N400 component reflects the process of accessing the meaning of the word (the activation of lexical/semantic representations) or that of integrating it with the prior context. Under the first view (lexical/semantic activation) (Federmeier and Kutas, 1999), the predictability-dependent N400 effect is often interpreted as facilitated lexical/semantic access to the predictable words owing to their successful pre-activation, and thus as evidence of prediction (Kutas and Federmeier, 2011; Lau et al., 2008). However, under the second view, the N400 effect could be explained in the absence of pre-activation. Following this, the reduced N400 amplitudes to predictable words could simply reflect how well the word happens to fit with the prior context upon perception (Brown and Hagoort, 1993; Hagoort et al., 2009). Of note, several recent studies support a multi-component view of the predictability-related N400 effects, with lexical/semantic predictability effects showing up earlier than integration effects (Brothers, 2015; Lau et al., 2016; Nieuwland et al., 2020).

Additional evidence supporting the connection between the N400 component and prediction arises from studies based on information theory (Hale, 2016; Levy, 2008; Smith and Levy, 2013), which assumes that listeners predict upcoming information by generating implicit, incremental probability distributions over possible continuations. Most notably, recent studies have found that N400 amplitudes correlate with the information-theoretical measure 'surprisal' (Frank et al., 2015; Heilbron et al., 2022; Michaelov et al., 2021). Word surprisal is the negative log probability of the word given the preceding sentence context and can be derived from language models, from simple n-gram models to large language models such as GPT-3. Moreover, word surprisal measures derived from sufficiently large models have been found to better model N400 amplitudes than human cloze probability judgments (Michaelov et al., 2021). Taking into consideration the principles of information theory and how word surprisal measures are mathematically formalized, this finding provides additional compelling evidence that language comprehenders depend on probabilistic knowledge for language processing (for a more detailed discussion, see Kuperberg and Jaeger, 2016).

In addition, word surprisal measures sit close to the concept of 'prediction error', a central component in probabilistic (Bayesian) computational models of language processing, which usually fall under the broader predictive coding framework (Clark, 2013; Friston, 2010). A key assumption in these models is that there is an asymmetry between the top-down and bottom-up information flow in the brain. The top-down flow carries predictions about incoming sensory data based on internal models that represent what the system already knows (prior distribution). These predictions are then compared to incoming sensory information (posterior distribution), and any discrepancies between the predicted and the actual input generate a bottom-up "prediction error" signal that is used to update the internal model to improve future predictions. Under this framework, N400 effects have been successfully modeled as semantic prediction errors (Rabovsky et al., 2018). In line with this, at the cognitive level, N400 effects of word predictability would reflect the additional retrieval of semantic features that have not been already pre-activated by the context (Kuperberg et al., 2020), fewer in the case of expected congruent than unexpected (either congruent or incongruent) words.

Another ERP component associated with the violation of expectations is the LAN —a left-anterior negativity also peaking about 400 ms after word onset—, which is typically encountered in response to unexpected morphological structure (Gunter et al., 2000; Koester et al., 2004; Söderström, Horne, and Roll, 2017). Predictive coding accounts consider it an indicator of morphological prediction error (Bornkessel-Schlesewsky and Schlesewsky, 2019). The LAN can be preceded by the earlier, word-category-sensitive ELAN—an early left-anterior negativity at 100–200 ms after word onset—which has been argued to represent prediction error for syntactic structure (Lau et al., 2006; Neville et al., 1991).

After the more specific initial negativities, the P600—a late positivity about 600 ms after word onset—responds more globally to the violation of *structural* expectations. The effect was originally reported for violated syntactic (Osterhout and Holcomb, 1992) and semantic (Kuperberg et al., 2003) structure. However, structural violations in other domains such as music also induce P600 effects (Besson and Faïta, 1995; Patel et al., 1998). A posterior and an anterior P600 have been differentiated (reviewed in van Petten & Luka, 2012). The posterior P600 is a response to syntactically/semantically implausible and incongruent words (i.e., syntactic and semantic structure anomalies) (Hagoort et al., 1993; Kuperberg et al., 2003), whereas the anterior P600 is observed for words that are syntactically/semantically unexpected, but still plausible within the context (Federmeier et al., 2007; Kuperberg et al., 2020). One interpretation is that the P600 is a member of the P300 family of ERP components but with a later peak due to the complexity involved in detecting violations in language and music (Coulson et al., 1998). The P300 complex is typically understood as having at least two subcomponents, the P3a, associated with novel stimuli, and the P3b, which is modulated by the probability of task-relevant stimuli (Squires et al., 1975). Particularly the P3b and the P600 seem to reflect similar mechanisms of conscious detection of incongruencies and subsequent updating of information, but see Frisch et al. (2003) and Osterhout (1999) for a different view.

Finally, an aspect to highlight is that the effects of word predictability generally appear earlier in the auditory than in the visual modality. In tasks using natural connected speech, an onset of negativities related to predictability has been observed even before the word recognition point of words (van Petten et al., 1999). Some interpret these effects as an early N400 effect, while others consider them to reflect pre-N400 effects related to the N200 component (with an onset around 150 ms), also called phonological mismatch (mapping) negativity (PMN) (Connolly and Phillips, 1994). This would index rapid and early detection based on phonological information that the word does not fit with the prediction. When it comes to memorized phrases (compared to unfamiliar phrases), such as in the case of proverbs, this signal can be even earlier (between 0 and 200 ms), which is attributed to highly specific pre-activation (Cermolacce et al., 2014).

Although N400, LAN, ELAN, and P600 modulations provide valuable information about prediction in language comprehension, given their post-target nature, this information is only partial about the full chain of processes involved in linguistic prediction. By definition, these should be at work earlier, that is, before the perception of the pre-activated input, and thus before the prediction is confirmed or disconfirmed. Following this rationale, several investigations have recently directed attention to ERP effects in the pre-target interval, which is more likely to reflect processes involved in the generation of the predictions. This research is preceded by a long history of examining pre-stimulus ERPs, which interestingly converge in showing that events that can be predicted are preceded by sustained event-related negativities.

3. Anticipatory negativities in cognition

Slow and sustained event-related negativities have long been a hallmark of anticipation in a variety of cognitive domains. Since the 1960 s, many studies have provided evidence that when one event signals that another event is about to occur in a temporally predictable manner, a sustained increase in negativity is observed in the electroencephalographic signal, typically over regions that are involved in processing the expected event. This electrophysiological phenomenon has been commonly linked to an internal state of anticipatory attention, or expectancy, towards the upcoming event.

Slow negativities are long-lasting deflections in the EEG signal—extending from hundreds of milliseconds up to several seconds—with a negative polarity relative to the stimulus baseline. They tend to emerge gradually in the fore period of predictable and relevant events, such as the execution of an instructed movement (Kornhuber and Deecke, 1965; Walter et al., 1964) or the presentation of a motivationally salient stimulus (Brunia and Damen, 1988). They have consequently been tightly linked to anticipatory mechanisms. These pre-target slow negativities typically surface over task-relevant regions (Khader et al., 2008; Rösler et al., 1997) and have been proposed to originate from synchronized excitatory postsynaptic potentials within cortical structures underneath their recording site (for an extensive discussion, see Birbaumer et al., 1990; Brunia et al., 2011). Accordingly, the elicitation of slow negativities at a particular electrode can be taken as a manifestation of an enhanced functional state in the underlying cortical area relative to other areas or other temporal intervals.

The *Contingent Negative Variation* (CNV) was one of the earliest demonstrations of a systematic slow negativity in humans. It was first observed when presenting participants with a warning signal (S1; a click sound) that announced the presentation of an imperative stimulus (S2; flashing lights) that required a motor response (Walter et al., 1964). A frontocentral slow negativity emerged gradually after S1, reached its peak before S2, and returned to baseline after the motor response. Importantly, the motor response was not a necessary condition to elicit the negativity (see also Ruchkin et al., 1986), but it was mandatory that a contingent relation between cue and target (S1 and S2) existed, as is captured by the name of the component.

The terminal phase of the CNV can be further decomposed into a motor-related component, the *Readiness Potential* (or Bereitschaftspotential), and a non-motor component, the *Stimulus Preceding Negativity* (SPN). In addition, the early and late components are underlied by another, longer-lasting CNV (sometimes referred to as “true” CNV), which has been related to timing processing (Macar and Vidal, 2004).

Early studies focusing on the SPN revealed that this component is sensitive to the value of the expected information (Damen and Brunia, 1985, 1987, 1994; Grünewald-Zuberbier et al., 1981; Kotani and Aihara, 1999). Specifically, the SPN precedes attended, imperative stimuli that provide informative feedback or that are motivationally salient, like performance feedback, monetary rewards, affective pictures, or painful aversive stimuli (van Boxtel and Böcker, 2004; for a review, see Böcker et al., 2001; Brunia, Hackley, et al., 2011; Poli et al., 2007). In fact, the subjective value of information is an important factor shaping the occurrence of the SPN, as was elegantly demonstrated by Donkers and van Boxtel (2005). In their study, participants performed a slot machine task in which they won money only when three consecutive images presented in a row were equal (e.g., XXX). Crucially, when the second image turned out to be different from the first (e.g., XY), the amplitude of the SPN decreased for the third and last image, as the participant could infer that it would not yield valuable information (i.e., a monetary reward). Note that even though the actual identity of the third and last image was unknown, the fact that it was not relevant to the subject’s goal anymore led to a relative decrease, or absence, of the SPN. Another study by Morís et al. (2013) reported similar findings. In an associative learning task, participants learned associations through trial and error and were given feedback on every trial. Critically, the amplitude of the feedback-related SPN decreased gradually as learning progressed, that is, as the feedback gradually lost its utility for participants because they had internalized the associations. Relatedly, Fuentemilla et al. (2013) found that the SPN was larger when participants were more uncertain about the upcoming feedback (i.e., monetary gain or not), and thus it was informative and relevant for them. These findings seem to suggest that processes other than anticipatory attention are likely to explain changes in SPN amplitude. More specifically, a recent proposal is that some of these processes might include the maintenance and adjustment of forward models (or “eligibility traces”) about expected targets (Bhangal et al., 2021; Ren et al., 2017).

In sum, there is abundant evidence of sustained negativities appearing prior to expected target stimuli. On the one hand, these sustained negativities have been functionally associated with the recruitment of attentional resources as a form of preparation for the bottom-up processing of the expected stimulus, as well as with time estimation mechanisms. Furthermore, a series of studies have recently suggested that the mechanisms underlying the SPN go beyond mere anticipatory attention and that they are intrinsically associated with the value of the information carried by the target.

4. Anticipatory negativities in language comprehension

Several studies in the field of language comprehension have reported slow and sustained pre-stimulus negativities that are compatible with anticipatory processing. The empirical strategy that has uncovered these sustained negativities is comparable to that of the cue-target paradigms adopted in studies of anticipatory processing in other cognitive domains—to analyze changes in brain activity before a target stimulus that can be anticipated from a preceding cue. Early examples displayed CNVs for words functioning as cues for other upcoming target word stimuli (Kutas et al., 2006). The task in those experiments was to judge the match of the words in terms of their phonological (Rugg, 1984a,b) or semantic properties (Butler et al., 1981; Rebert and Lowe, 1980). The CNV in these linguistic contexts was markedly left-lateralized as compared to the CNV found for spatial or visual feature matching of face identity-matchin g (Bentin et al., 1985). However, studies on prediction in more naturalistic language processing have been scarcer.

In fact, cues in language are not always easy to condense and discretize. They can consist of something as simple as a single word or as complex, composite, and gradual as a sentence context with all its associated features (morphological, phonological, syntactic, semantic, and/or pragmatic). Furthermore, the cue and the target can be two distinct linguistic items, or they can be two parts within the same item (Huettig et al., 2022). In this regard, some more recent studies have focused on prediction occurring in more naturalistic comprehension, either within sentences (sentence contexts serving as cues to pre-activate words embedded within or finalizing them) or within words (the first portion of a word serving as a cue to pre-activate its last part). In both cases, anticipatory negativities have been found to develop between the cue and the target, with amplitudes that are modulated by the predictive value of the cues. Given the conditions under which they arise, and as we will discuss in the following sections, these anticipatory signals are likely associated with the prediction of upcoming linguistic material and we will refer to them as prediction negativities.

4.1. Using sentence contexts to predict

Sentence comprehension involves at least two core components: accessing the meanings of words and integrating them to form a cohesive, internal representation of the meaning of the context. This internal representation of the sentence context contains higher-level information that goes beyond what is conveyed by single words or features, and that exerts a strong top-down influence on how upcoming linguistic input is processed (van Petten, 1995). Furthermore, sentence contexts can have different predictive strengths about upcoming words. Specifically, sentence contexts that are highly constraining (HC) lead strongly to a single word (e.g., “Don’t touch the wet” leads strongly to the word “paint”), whereas low constraining (LC) sentence contexts admit multiple continuations (e.g., “There was nothing wrong with the” can be followed either by "car", "kid", or "job"). Therefore, HC contexts are more likely to lead (or lead more strongly to) the pre-activation of specific representations. Under this assumption, researchers have zoomed into the brain activity that precede target words embedded in HC and LC sentence contexts, aiming to detect neural markers of prediction in sentence comprehension. Applying this strategy, several recent studies have revealed prediction negativities, that is, anticipatory sustained negativities preceding sentence-final words in HC (relative to LC) contexts (Grisoni et al., 2017, 2021; León-Cabrera et al., 2017; 2019; 2021; Li et al., 2017) (Fig. 1).

For example, in a study by Li et al. (2017), participants read HC and LC sentences that contained a verb that provided critical information to predict an upcoming target “To practice calligraphy, my brother bought brand-name *brush pens* and took them home.” (critical verb underlined and target word in italics). In HC sentence contexts (relative to LC), a prediction negativity was triggered by the verb and encompassed the 1400 msec leading up to the target word, with maximal amplitudes over left frontal electrodes (Fig. 1**A**). They interpreted this prediction negativity as reflecting increased processing costs to generate/maintain the pre-activated representations in HC contexts.

Other studies have reported prediction negativities when introducing a noticeable delay (up to 1–2 seconds) between the sentence context and the target word, thus more closely emulating CNV and SPN paradigms. An advantage of increasing the pre-target interval is that it ensures that comprehenders have sufficient time to pre-activate not only semantic but also more specific information including the word-form (Freunberger and Roehm, 2016; Ito et al., 2016; although word-form pre-activation might also take place at a normal reading pace, DeLong et al., 2021). Therefore, the introduction of a delay is an empirically

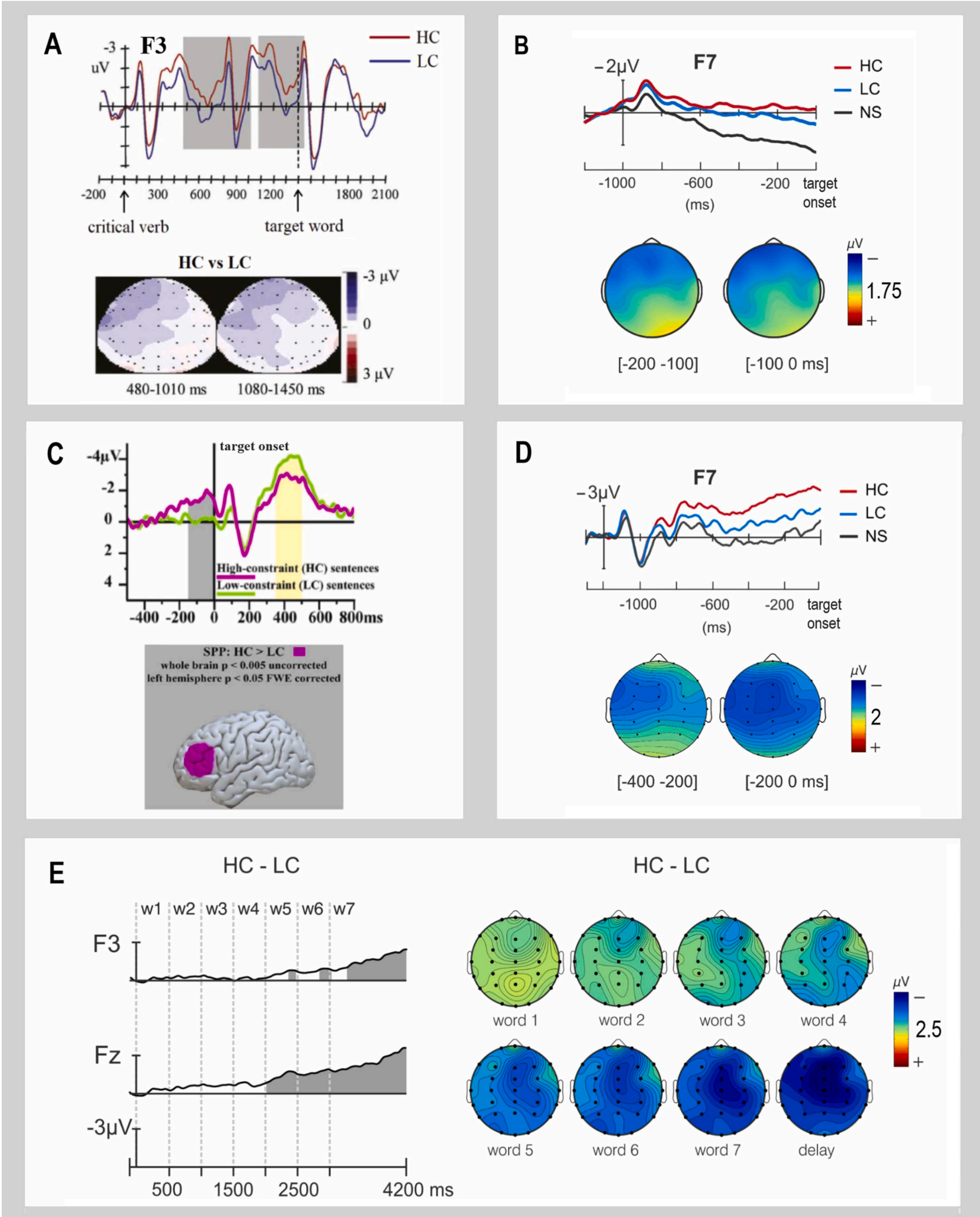

Fig. 1. A) Adapted from Li et al. (2017): left anterior negativity triggered by critical verbs that were strongly constraining of an upcoming target word (HC relative to LC contexts) in written comprehension. B) León-Cabrera et al. (2017): left anteriorly distributed anticipatory negativity in the 1-second interval between HC sentence contexts and their sentence-final (target) words (relative to LC and NS) in speech comprehension.The topographies show the difference between HC and LC. C) Adapted from Grisoni et al. (2021): anticipatory negativity preceding sentence-final (target) words in HC (relative to LC) sentences in speech comprehension, source-localized in the left inferior frontal cortex. D) León-Cabrera et al. (2019): left anteriorly distributed anticipatory negativity in the 1-second interval between HC sentence contexts and their sentence-final (target) words (relative to LC and NS) in written comprehension. The topographies show the difference between HC and LC. E) León-Cabrera et al. (2019): ultraslow negativity (<5 Hz) emerged over frontocentral sites already at the fourth to fifth word of the sentences for HC (relative to LC and NS) sentences.

valuable strategy to boost prediction at low representational levels. This can facilitate the identification of brain indices associated with pre-activation that might be later investigated in more naturalistic settings.

An early instance of a study adopting this approach is the one by Besson et al. (1997), who investigated ERP responses during a pause between a sentence context and its final word. They presented HC (proverbs) and LC sentence contexts, either visually or auditorily, and introduced a 600 ms pause before the final word in half of the trials. In the visual modality, a slow negative potential developed during the pause, with larger amplitudes in LC than HC contexts, interpreted as a CNV reflecting greater expectancy in LC contexts. In contrast, in the auditory modality, they observed a marked emitted potential (instead of slow component) with more negative amplitudes were observed in LC compared to HC contexts. The effect was interpreted as reflecting surprise due to the unexpected interruption of fluent speech, greater in HC (relative to LC) contexts, where very strong expectations about the continuation were suddenly unfulfilled due to the pause.

Inspired by these findings, León-Cabrera et al. (2017) conducted a speech comprehension study using a similar delay paradigm, but with some adjustments. In the experiment, sentences were presented with varying levels of contextual constraint (HC and LC), and a control condition was included where predicting the final word was not possible at all because the sentences were non-semantic (NS) —meaningless sentences created by randomly swapping the sentence's vowels. In addition, to eliminate the effect of surprise, the pre-target pause (1000 ms) was present in all trials. In line with the results from Besson et al. (1997), the ERP analysis of the pre-target interval revealed the development of an anticipatory negative potential. However, in this case, it was larger (more negative) at increasing levels of contextual constraint (HC > LC > NS) and maximal at left anterior sites (most prominent at F7 electrode) (Fig. 1**B**), in line with the prediction negativity in Li et al. (2017). One task difference that could explain this striking divergence is that Besson et al. used familiar proverbs, as opposed to unfamiliar HC sentences. As fixed expressions in memory, proverbs may involve different predictive mechanisms than familiar sentences, relying on categorical expectations instead of context-based, probabilistic expectations (Vespignani et al., 2010), which might lead to distinct brain responses (Cermolacce et al., 2014; Proverbio et al., 2009). Therefore, the prediction negativities in León-Cabrera et al. (2017) and Li et al. (2017) might capture mechanisms associated with context-driven probabilistic prediction, thus reflecting the strength and/or accuracy of the prediction as a function of constraint. In a later study, the authors adapted the same task to the visual modality (León-Cabrera et al., 2019) (Fig. 1**D**). Participants read sentences (HC, LC, or NS) that were presented one word at a time, and that also contained a 1-second delay before the sentence-final word. In line with the findings in speech comprehension, a left-anterior prediction negativity developed in the delay with larger amplitudes for HC (than LC and NS), suggesting that the mechanism underlying this prediction negativity is not entirely dependent on input sensory modality.

More recently, adopting a similar approach, Grisoni et al. (2021) observed a prediction negativity in a speech comprehension task (Fig. 1**C**). Specifically, they also reported the emergence of a frontally-distributed prediction negativity in the 1100 ms delay preceding final words in HC (relative to LC) sentences. An important contribution of this study is that they located the main source of the prediction negativity in the left inferior prefrontal cortex. In addition, their sentence contexts were designed to constrain to either animal-related or tool-related concepts. Within the semantic categories, they also found a prediction negativity, but additionally detected category-specific sources in visual-related and motor-related areas for animals and tools, respectively. In fact, an earlier study by Grisoni and colleagues (2017) had already shown content-specific prediction negativities, emerging over dorsolateral motor regions when it preceded a hand-related verb ("write") and over ventral motor regions when it preceded a face-related verb ("talk"). The somatotopy motivated the authors to interpret their prediction negativity as capturing the "semantic features of the anticipated stimulus" (Pulvermüller and Grisoni, 2020).

A characteristic feature of all these prediction negativities is that they build up over time before the presentation of the target. Indeed, in the aforementioned study by León-Cabrera et al. (2019), the prediction negativity preceded sentence-final words by several seconds, suggesting that the underlying anticipatory process can begin substantially early during sentence comprehension. Specifically, they found that an ultraslow negativity (low-pass-filtered at 5 Hz) emerged over frontocentral sites already at the fourth to fifth word of the sentences for HC (relative to LC and NS) sentences (Fig. 1**E**). Previous studies in the field of sentence processing had reported ultraslow negativities (encompassing many words). These negativities are likely associated with longer-lived cognitive processes involved in sentence comprehension, which underlie co-occurring, shorter-lived processes at the word or sub-word level, like, for example, the N400 component (Kutas and King, 1996). Likewise, the results in León-Cabrera et al. (2019) suggested that there was a longer-lived process, as reflected in the broadly distributed ultraslow prediction negativity, underlying other, shorter-lived processes, like the more spatiotemporally constrained, left-anterior prediction negativity that emerged in the last milliseconds before the presentation of the expected target word. Previous studies have consistently tied cross-clause negativities to increased working memory (WM) demands when processing sentences with difficult syntactic or conceptual configurations (Fiebach et al., 2002; Kutas and King, 1996; Matzke et al., 2002; Münte et al., 1998; King and Kutas, 1995). Following this, the authors argued that this ultraslow sentence-level modulation could similarly capture increased WM costs, in this case resulting from top-down prediction in HC sentence contexts.

One critical aspect of many of the reviewed studies is that half of the targets in every condition (HC and LC) were semantically incongruent (Besson et al., 1997; León-Cabrera et al., 2017; 2019; 2021; Li et al., 2017; Grisoni et al., 2017). This is usually done with the goal of assessing word predictability effects on the N400 component—that is, that incongruent and unexpected targets, which could not be predicted by the participant, elicit larger N400 amplitudes than the congruent and predictable ones (reviewed in Kutas and Federmeier, 2011). Consistently, all these studies report N400 congruency effects, which could be interpreted as reflecting post-target prediction match or mismatch (as discussed in Section 2). Notably, Grisoni et al. (2021) did not include incongruent or highly unexpected words but target nouns differed in predictability (cloze probability). In the same vein as the studies manipulating congruency, amplitudes were more negative in the HC (than in the LC) condition. Further, the authors reported an inverse correlation between pre-target prediction negativities and post-target N400 amplitudes: more constraining contexts resulted in increased negativities *before* target words and reduced N400 negativities upon hearing the target words. Notably, pre-target prediction potentials did not correlate with N200 amplitudes, strengthening the idea that prediction negativities may indeed be functionally related to lexical/semantic pre-activation.

Thus, the combined pre- and post-target ERPs in all these studies are consistent with prediction. However, each task might lead to the pre-activation of different types of features (e.g., only semantic or also orthographical). The reading comprehension task by León Cabrera et al. (2019), for instance, showed that incongruent words in HC contexts triggered a larger negativity than congruent ones as early as 200–250 ms post word onset. In the visual domain, N250 responses decrease for repeated words (Holcomb and Grainger, 2006), indicating facilitated orthographic processing (Morris et al., 2008). Hence, the early negativity strengthens the idea that participants were generating highly specific predictions, perhaps even at the level of word form, which might have been enabled by the relatively slow rate of presentation of the words in this task (Freunberger and Roehm, 2016; Ito et al., 2016).

Relatedly, the fact that these paradigms incorporated salient delays might have also critically boosted prediction effects by providing more time than usual to form predictions, and/or by inducing a strategic increase of attention towards the target word. Although the presence of delays may have influenced the kind of cognitive processes engaged during the task, it should be noted that the sustained negativities found in these paradigms are similar to those in studies that do not incorporate delays (Li et al., 2017). Likewise, the presence of incongruent words (which are not a common phenomenon in naturalistic language settings) does not seem to play a determinant role in the elicitation of these anticipatory negativities, given that they also emerge when only congruent words are included (Grisoni et al., 2021). Another argument against these task-related factors—the presence of delays and/or incongruent words—driving the emergence of prediction negativities in these tasks is that they are a constant variable in both levels of constraint (HC and LC) and, therefore, cannot explain the differences between the two conditions. Nevertheless, future studies should directly test the influence of these factors on the elicitation of the prediction negativities. For example, a study could manipulate the length and quality of the silent delays to investigate to what extent the anticipatory sustained negativities are conditioned by the pauses.

4.2. Using phonological cues to predict

In parallel to the aforementioned work, several studies have reported strikingly similar effects in tasks employing specific phonological cues rather than sentence contexts to study prediction. The advantage of the phonological cues is that ERPs can be time-locked to cue onset, which allows researchers to study the process of prediction from its very beginning. There is, by now, extensive evidence that individuals use phonological information to generate predictions. Specifically, listeners have been observed to take advantage of prosodic (Hjortdal et al., 2022; Lozano-Argüelles et al., 2020; Roll et al., 2013; Sagarra and Casillas, 2018; Söderström et al., 2018) and segmental information (Allopenna et al., 1998; Dahan et al., 2001) and even fine-grained phonetic details (Archibald and Joanisse, 2011; Beddor et al., 2013; Salverda et al., 2014) to pre-activate anticipated upcoming speech. In ERP studies, more constraining phonological cues typically elicit an increased negativity starting before the cued target forms appear in their entirety in the speech input (Dufour et al., 2013; Gosselke Berthelsen et al., 2018; Hjortdal et al., 2024; Hunter, 2013; Roll et al., 2010, 2015, 2017; Söderström et al., 2016; Söderström, Horne, and Roll, 2017; Söderström, Horne, Mannfolk, et al., 2017).

One line of research has investigated the predictive function of phonological word accent cues in Swedish and Danish. Swedish words bear tonal information that is conditioned by following suffixes (Bruce, 1977). The noun ^{1}*hatt* ‘hat’, for instance, is realised with a low tone (accent 1). The tone is retained in the presence of the definite singular suffix *-en* in ^{1}*hatt-en* ‘the hat’ (literally ‘^{1}hat-the’) but is replaced by a high tone (accent 2) when the stem is combined with the indefinite plural suffix *-ar* in ^{2}*hatt-ar* ‘^{2}hat-s’. The Danish variety of word accents distinguishes between a creaky voice quality, *stød*, and a modal voice quality, *non-stød* on the stem, but there are similar associations between prosody and morphology (Fischer-Jørgensen, 1989). Accent 1 is typically more constraining than accent 2 as it has a smaller set of possible endings (Hjortdal et al., 2022, 2024; Roll, 2022; Söderström et al., 2016). Researchers exploited this intrinsic asymmetry in the degree of constraint of phonological cues to study neural activity timed directly to the phonological cue, allowing them to study prediction unfolding from the specific moment that the target is cued. They have typically embedded speech fragments with phonological cues in identical (and therefore equally constraining) contexts. For the Swedish word accent example above, this could be a sentence like ‘For Christmas, I got ^{1}hat-the / ^{2}hat-s.’ In these types of neutral contexts, the phonological cues (the different tones on the stem, marked with superscript numbers) are then virtually the only features that make a difference in constraint and, therefore, any divergence in predictive activity can be directly related back to them.

The more constraining phonological cue elicits a stronger negativity in the ERP signal. This negativity has been labeled PrAN (pre-activation negativity) and differs from the previously described anticipatory negativity for sentence contexts with respect to how the ERPs are time-locked. The PrAN is time-locked directly to the onset of the predictive cue and can therefore track the pre-activation of the predicted ending from its beginning. This is impossible in the case of sentence contexts where predictions build up incrementally over the course of the sentence culminating in the pre-target time window (see Section 4.1.). The PrAN typically starts as early as 136 ms after phonological cue onset, and the negativity can last for well beyond 400 ms (Fig. 2**A, B**). While the negative deflection is overall most pronounced at anterior and central electrodes, the PrAN has been argued to comprise two phases that might have partly distinct neural substrates (Fig. 2**D**; Roll et al., 2023). The first phase—before 200 ms—has a left posterior topographical distribution (Roll et al., 2015; Söderström, Horne, Mannfolk, et al., 2017), and correlations with blood-oxygen-level-dependent (BOLD) effects have suggested sources mostly in and around the auditory cortices of the left temporal lobe (Roll et al., 2015; Söderström, Horne, Mannfolk, et al., 2017). In contrast, the second phase—after 200 ms—typically has an anterior (Gosselke Berthelsen et al., 2018; Hjortdal et al., 2022; Novén, 2021; Roll et al., 2010) or left anterior distribution (Roll et al., 2015; Söderström, Horne, and Roll, 2017; Söderström, Horne, Mannfolk, et al., 2017) with BOLD-effect correlations mainly in the left inferior frontal gyrus (Roll et al., 2015, 2017; Söderström, Horne, Mannfolk, et al., 2017; Söderström et al., 2018). The BOLD response differs slightly depending on what type of linguistic material can be predicted. When the prediction is syntactic rather than morphological, for instance, an additional generator has been found in the left anterior insula (Söderström et al., 2018), which is associated with syntactic predictions (Jakuszeit et al., 2013) (see Fig. 2**C**). Topographical comparisons and proposed sources suggest that the second phase of the PrAN resembles the prediction negativity found in the pre-target time windows in the sentence context paradigms discussed in Section 4.1. The effect is functional rather than acoustically conditioned since it was replicated in a dialect where the tones represent an acoustic mirror image as compared to the previously studied Central Swedish (Roll, 2015). The PrAN has, likewise, been dissociated from the N1 component, obtained for acoustically salient features. Specifically, Roll et al. (2013) observed a PrAN overlapping with the P2 component for the highly constraining but acoustically non-salient low accent-1 tone. When the words were deprived of their lexical content, no PrAN was found, but instead an N1 increase in the preceding time window for the acoustically salient high accent 2 tone.

Aligned with the majority of studies investigating prediction, the PrAN has been elicited in paradigms that include prediction violations where phonological cues are followed by uncued targets such as incongruent suffixes. Such prediction violations lead to increased reaction times (Söderström et al., 2012; Clausen and Kristensen, 2015) as well as the well-known post-target responses related to prediction error reviewed in Section 2. Depending on the experimental task and the type of cued information, encountering an uncued target either elicits an increased N400 (Gosselke Berthelsen et al., 2018; Hjortdal et al., 2022) or a LAN (Novén, 2021; Söderström, Horne and Roll, 2017) followed by a P600 response (Gosselke Berthelsen et al., 2018; Hjortdal et al., 2022; Novén, 2021; Roll, 2015; Roll et al., 2010, 2013, 2015). The emergence of the traditional markers of failed prediction in addition to the pre-target PrAN further strengthens the claim that listeners use the associations between prosodic cues on word stems and target suffixes to pre-activate the endings before they are realized in the speech signal.

Building on the robust association between PrAN and predictive strength, researchers have explored the effect in other contexts where the phonological cues differ in how constraining they are. It has long been a fundamental assumption of many models of spoken word recognition that listeners incrementally pre-activate potential word candidates in parallel as soon as the first few speech sounds of a word are available (Marslen-Wilson, 1987; McClelland, 2013; Norris et al., 2016). In line with this, the PrAN is modulated by the lexical constraint provided by the first few speech sounds in a word. The more constraining the word beginning, the more negative the PrAN (Hjortdal et al., 2024; Söderström et al., 2016; Roll et al., 2017). The constraint can be estimated from the number of complete words (and their respective probabilities) that can be formed from the word beginning: the fewer lexical competitors, the more constraining the word beginning. For example, the word onset *ts-* can form relatively few complete words, like *tsunami*,

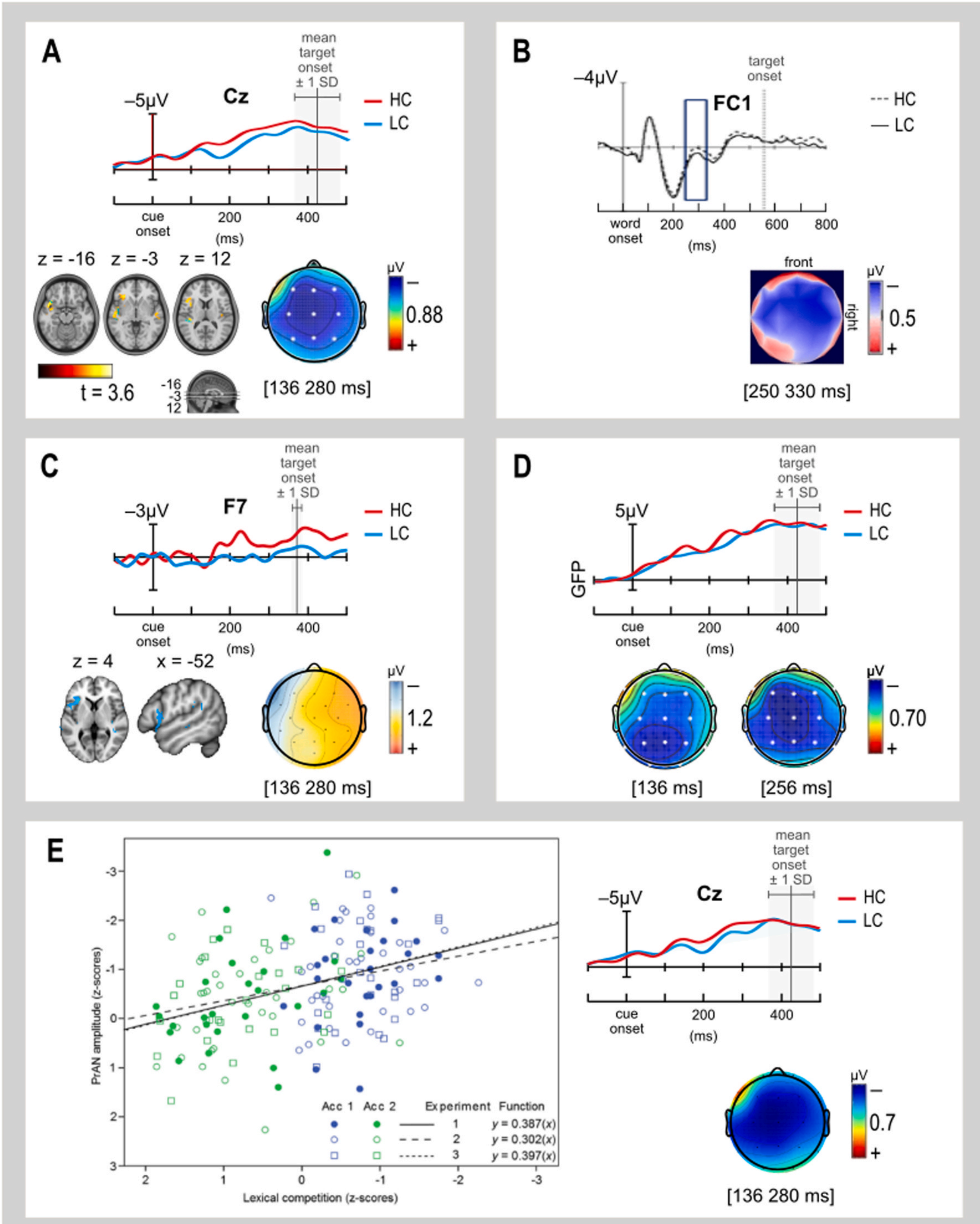

Fig. 2. A) Roll et al., (2015): Pre-ativation negativity (PrAN) for tonal word onsets cueing suffixes and correlated BOLD contrast. B) Dufour et al., (2013): Negativity for word onsets cueing endings with different neighborhood density. C) Söderström et al., (2018): PrAN for boundary tones cueing syntactic structure and correlated BOLD contrast. D) Roll et al., (2015): Global field power (GFP) peaks indicating bi-phasic activity. E) Söderström et al., (2016): PrAN for word onsets cueing endings with a high and low number of possible continuations; correlation between PrAN and lexical competition. HC = high constraint; LC = low constraint. Note that in the PrAN experiments, ERPs are locked to cue onset. Target onset varies. We illustrate mean target onset with gray lines, while the gray-shaded area shows the variation (+/- one standard deviation). All figures have been adapted slightly from the original papers for a more coherent presentation.

whereas *st-* can have many completions, including *start* and *stop*. Therefore, *ts-* is more constraining than *-st*. More negative amplitudes have been reported for words with few lexical competitors in an early time window (136–450 ms), not only for Swedish (Roll et al., 2017; Söderström et al., 2016) but also for English and French (Dufour et al., 2013; Hunter, 2013). Early works interpreted similar effects as increased positivities for the condition involving larger numbers of possible continuations and higher lexical competition (Hunter, 2013; Roll et al., 2010, 2013). However, global field power (GFP) measures of ERPs and their correlation with an increased hemodynamic signal showed that the polarity was rather a negativity for HC phonological cues (Roll et al., 2015, 2017; Söderström et al., 2016). The GFP quantifies the amount of activity of the electrical field across all electrodes. GFP peaks were originally proposed for detecting ERP component maxima (Lehmann and Skrandies, 1980). Based on pronounced GFP peaks, the PrAN studies argued that HC phonological cues rather than LC cues drive the effect (Roll et al., 2015; Söderström et al., 2018).

A gap in the existing research are studies investigating the effects of differently constraining phonological cues in neutral sentences–but with a 1000-ms delay preceding the target. Delays within words would indeed sound unnatural, but such a study is possible when the phonological cue is realized on a separate word. English definite or indefinite articles are examples of such words. While the articles *a* and *the* [θə] precede words beginning with consonants, *an* and *the* [θi:] precede vowels. Our intuition is that the latter are more constraining and thus stronger phonological cues because they signal vowels, a more limited group of speech sounds, whereas *a* and *the* [θə] are followed by a larger class of consonants, but this should of course be backed up by a corpus study. Similar to the studies of Leon-Cabrera et al. (2017, 2019) and Grisoni et al. (2017), (2021), delays could be imposed between articles and following words. We hypothesize that a stronger negative potential would build up following *an* and *the* [θi:] than *a* and *the* [θə]. A crescendo of the enhanced negativity prior to the target noun, similar to what has been reported by León-Cabrera and colleagues, would further support that the negativity reflects anticipation rather than processing of the phonological cue itself.

4.3. Potential mechanisms underlying anticipatory negativities in language comprehension

A close inspection of both bodies of work (i.e., using sentence contexts or phonological cues) reveals the common finding of the development of prediction negativities —anticipatory sustained negativities in response to strongly predictive cues that allow the pre-activation of upcoming linguistic input at different representational levels. These cues range from simple and discrete phonological information (prediction negativities for phonological cues, or PrAN), to the more complex and diverse information embedded in sentence contexts (prediction negativities in sentence contexts). Most relevantly, the latency and morphology (wave shape) of these components is strongly consistent with anticipatory processing: they not only precede the target, but ramp up before its presentation, similar to the anticipatory negative potentials reported in other cognitive domains and in simpler linguistic tasks (see Sections 3 and 4). Their topographical distribution is more heterogeneous, but there is a common left anterior locus across prediction negativities for sentence contexts and phonological cues. Next, we will discuss their similarities and differences and interpret them in the light of other, well-established ERPs associated with predictive language processing and anticipatory processing.

The anticipatory character of prediction negativities is supported by the fact that their amplitude is modulated by how strongly linguistic cues constrain the possible targets. As mentioned above, prediction negativities have similar antecedent conditions in that they are both modulated by constraint, either from a preceding sentence context or a phonological cue. For example, the amplitude of the PrAN (i.e., prediction negativities for phonological cues) varies continuously with different levels of constraint. Recently, Hjortdal et al. (2024) reanalysed data from three previous experiments which investigated the predictive value of word accents (Roll, 2015; Novén, 2021; Hjortdal et al., 2022). Using combined pronunciation lexica and frequency lists, they calculated measures of cohort entropy, an estimate of the uncertainty about the lexical identity of unfolding words, upon hearing word beginnings up until and including word accents. Entropy can be understood as the average or expected surprisal of an outcome: if many lexical candidates are equally likely, entropy is high whereas if one candidate has high probability, entropy is low. Noticeably, Hjortdal et al. reported a correlation between brain potentials and cohort entropy (while controlling for phoneme surprisal). In other words, the more constraining the first few phonemes of a word, the more negative the ERP amplitude. This observation is similar to what has been reported in the non-linguistic CNV literature (Bennett et al., 2015; Gómez et al., 2019). For instance, Gómez et al. (2019) found that when participants had stronger prior expectations of cue validity in a Central Cue Posner Paradigm, ERP amplitudes preceding the target stimulus were more negative. In a similar vein, the amplitude of the prediction negativities in sentence contexts seem to follow a linear trend across HC, LC, and non-semantic sentence contexts (León-Cabrera et al., 2017, 2019) (Figs. 1**B,** 1**D**). This relationship could be further investigated by also incorporating conditions of medium constraint or a continuous range of constraint levels.

Another common feature of the prediction negativities is that they precede the target and last until its encounter. However, they can differ in duration. The PrAN (i.e., the prediction negativity for phonological cues) is time-locked to the onset of the predictive cue (e.g., a tonal change) and it has a relatively short duration (about 400 ms), emerging and resolving within words (Fig. 2). The specificity of the cue, critically, allows the experimenter to isolate the time window when the pre-activation—triggered by the predictive cue—is likely to initiate. This differs from the studies in which the ERPs are time-locked to the onset of a delay between the sentence context and the final word (Grisoni et al., 2017; 2021; León Cabrera et al., 2017; 2019) (Figs. 1**B,** 1**C,** 1**D**). In these cases, longer-lasting predicion negativities build up in the delay (for about 600–800 ms) until the presentation of the sentence-final word, and are morphologically more similar to the SPN, which has often been associated with increased anticipatory attention to relevant stimuli (see Section 3). Therefore, one possibility is that these longer-lasting prediction negativities also reflect an increased level of attention and expectancy to the presentation of the incoming word. Interestingly, attention might be correlated with the degree of specificity of the prediction. Under a predictive coding framework, modulations in the level of attention would be associated with the expected degree of precision of the predictions that are being generated (Friston et al., 2018; Walsh et al., 2020). As such, situations where individuals make more precise predictions (e.g., after a strongly constraining cue) might go hand in hand with a state of greater anticipatory attention to the incoming bottom-up information. Therefore, the amplitude of the anticipatory negativities might reflect an increased state of expectation that is functionally related to the specificity of the prediction. The previous idea also fits well with existing evidence in the oscillatory domain showing that alpha (8–12 Hz) and/or beta (~13–25 Hz) power is decreased in the anticipatory interval of words in HC relative to LC sentence contexts (León-Cabrera et al., 2022; Li et al., 2017; Rommers et al., 2017; Terporten et al., 2019; Wang, Hagoort, and Jensen, 2018), that is, under the same conditions that elicit prediction negativities. Interestingly, alpha/beta power decreases have been associated with optimal states for the processing and representation of information (Hanslmayr et al., 2012). For example, a recent combined EEG-fMRI study found that alpha/beta power decreases correlated with brain patterns containing more specific information about stimuli (videos or melodies) during their perception and recall (Griffiths et al., 2019). These findings were interpreted as indicative that alpha/beta power decreases provide favorable conditions for accurately representing information. Following this rationale, the alpha/beta power decreases preceding sentence-final words in HC contexts might similarly reflect the presence of a more rigorous and well-specified internal representation (relative to LC contexts) associated with the pre-activation of stimulus-related low-level features.

On the other hand, as previously mentioned, the shorter-lived and more specific nature of the PrAN (lasting about 400 ms on average) might be more strongly tied to the moment in which the pre-activation is initiated. Of note, the second phase of the PrAN shows a frontal, left-lateralized distribution (Roll et al., 2015, 2017; Söderström et al., 2016, 2018) (Figs. 2A, C, and D), which strikingly converges with the observation that prediction negativities in sentence contexts are largest at left frontal electrodes (i.e., F7 electrode) in the last milliseconds immediately preceding the target (Grisoni et al., 2021; León-Cabrera et al., 2017, 2019; Li et al., 2017) (Figs. 1A, B, C, and 1D). A possibility is that both types of prediction negativities reflect the same process. In line with this notion, source-localization findings suggest that, in both cases, the effect stems from activity in the left inferior frontal cortex (LIFC) (Grisoni et al., 2021; Roll et al., 2015, 2017; Söderström et al., 2018). The LIFC is involved in multiple language-specific (Friederici, 2012; Hagoort, 2017) as well as domain-general functions (Fedorenko and Blank, 2020; Novick et al., 2010). However, one of its roles that might be particularly compatible with the elicitation of the left anterior prediction negativities is that of selection and retrieval of linguistic representations (Matchin et al., 2017; Thompson-Schill et al., 1997; Zhuang et al., 2014). More specifically, during sentence comprehension, the anterior and posterior LIFC would utilize context-based information to mediate the controlled selection and retrieval of linguistic representations, respectively (Lau et al., 2008). As such, the LIFC might be a key region supporting the generation of context-based predictions (Silcox et al., 2023), and left frontal prediction negativities immediately preceding targets may be associated with this process, both in response to phonological cues and to sentence contexts. In this regard, it is interesting to note that the sources of the PrAN vary as a function of whether the prediction is about morphological or syntactic information (Roll et al., 2015; Söderström et al., 2018). Likewise, prediction negativities in sentence contexts can have topographical distributions and sources that are related to the semantic category of the expected words (Grisoni et al., 2017, 2021). These findings are suggestive that the topographical distribution and/or the underlying sources of prediction negativities may serve as online indices of pre-activation at different levels of representation (e.g, morphological, syntactic, or semantic).

An unresolved question is whether other types of linguistic cues also elicit prediction negativities. In this regard, a few ERP studies have looked into the role of grammatical cues. For instance, van Petten and Kutas (1991) found that function words (e.g., determiners, pronouns, prepositions, etc.) elicited a larger negativity than content words (e.g., nouns, verbs, adjectives) about 400–700 ms post-word onset. The authors interpreted the negativity as a type of CNV and argued that function words cued the forthcoming presentation of a more informative, content word. In a more recent ERP study, Huang et al. (2023) investigated the effects of Mandarin Chinese animate- and inanimate-constraining classifiers as cues to following nouns. In both conditions, a negativity built up in a 1000-ms interstimulus window preceding the target nouns. These studies hint at grammatical cues like gender-marked articles and classifiers being used to predict upcoming words. To further attest this, future studies could quantify and directly manipulate the degree of constraint of these grammatical cues. We expect that prediction negativities will arise time-locked to the onset of the grammatical cues, i.e., with larger amplitudes for greater levels of constraint.

Overall, prediction negativities might serve as indices that a certain linguistic cue (e.g., sentence context or phonological cue) has triggered the pre-activation of a specific representation. One relevant contribution of the characterization of these prediction negativities is that, given their anticipatory nature, they can contribute to a more nuanced understanding of the cognitive processes involved in linguistic prediction. In contrast to post-target ERP components, like the N400, LAN, or P600, prediction negativities emerge in the pre-target interval and therefore are more likely to directly relate to the pre-activation process, rather than its aftermath. Therefore, prediction negativities are bound to provide complementary information about the different phases and operations within the predictive chain (see Table 1). As an example, under predictive coding frameworks, pre-target prediction negativities would reflect top-down pre-activation, while post-target ERP indices would capture bottom-up prediction error and model updating. The shorter-lived PrAN may be more adept at pinpointing the precise moment that the pre-activation of a specific representation happens, and it can be combined with relatively punctate post-target ERPs, such as the N400, to

Table 1
Overview of ERP components that have been related to linguistic prediction. The prediction negativities for sentence contexts and phonological cues (PrAN) are pre-target components that precede a predictable target word. They are more likely to capture processes more directly involved in pre-activating representations. On the contrary, the left anterior negativity (LAN), N400, and P600 are post-target components which have been related to processing costs.

Stage	ERP component	Latency	Amplitude modulation	Topographical distribution	Associated process
Pre-target	Prediction negativity for sentence contexts	Builds up for several hundreds of milliseconds before the target	Larger for sentence contexts that are strongly predictive of the target	Typically left frontal, but it can be content-specific	Pre-activation mechanisms
	Pre-activation negativity (PrAN)	The onset is 136–280 ms following phonological cues and it can last for hundreds of milliseconds	Larger for phonological cues that are strongly predictive of the target	Left frontal and frontocentral	Pre-activation mechanisms
Post-target	N400	Peaks 400 ms after target onset	Larger for targets that violate semantic expectations	Centro-parietal	Match or mismatch of semantic predictions
	LAN	Peaks 400 ms after target onset	Larger for targets that violate morphological expectations	Left frontal	Match or mismatch of morphological predictions
	P600	Peaks 600 ms after target onset	Larger for targets that violate structural expectations	Anterior or posterior	Detection and repair of structural violations

simultaneously examine pre-activation and its facilitatory effect on bottom-up processing. On the other hand, the longer-lived prediction negativities appearing between sentence contexts and sentence-final words might additionally capture states of increased attention towards the expected input. We have put forward the hypothesis that both the shorter-lived and longer-lived prediction negativities that share a left frontal distribution might reflect the context-driven selection and/or retrieval of expected representations, although future studies are necessary to directly test this hypothesis. Likewise, more research is needed to understand to what extent these prediction negativities could be informative of the type of content that is being pre-activated.

Finally, future experiments could simultaneously track the development of different types of prediction negativities within the same sentence. For example, within sentences, certain words might provide key information to predict another, incoming word. This could be indexed by left frontal prediction negativities developing between the two items, as in the study by Li et al., (2017) (Fig. 1A). On the other hand, as a sentence unfolds, comprehenders gradually build a high-level representation of the meaning of context, which may be constantly used to pre-activate information at multiple representational levels (e.g., Kuperberg and Jaeger, 2016). Interestingly, the effects of this continuous top-down issuing of predictions (at different representational levels) might be imprinted in slower and broadly distributed prediction negativities building up over the course of sentence processing (León-Cabrera et al., 2019) (Fig. 1E). These slower modulations might therefore capture longer-lived cognitive operations underlying the other, shorter-lived and more specific prediction negativities, like the PrAN, which, as previously discussed, might index when a more specific linguistic representation (e.g., morphological, syntactic, semantic, or other) has been pre-activated. Investigating both shorter- and longer-lasting prediction negativities in the same task could be enlightening about which information is being used to generate predictions in a particular situation.

5. Concluding remarks

There is a consensus that language users predict linguistic input during comprehension, at least under certain circumstances. Most ERPs used to measure prediction in language primarily focus on processes triggered after perceiving pre-activated words, while there is limited knowledge regarding ERP signatures preceding them. In this pursuit, recent studies have converged in showing anticipatory negativities preceding words or word segments that can be predicted from sentence contexts or phonological cues. Based on their commonalities in terms of functional sensitivity and spatiotemporal features, as reviewed here, we propose that these negativities may reflect cognitive operations involved in pre-activating linguistic input at different representational levels. Crucially, these prediction negativities have attributes that are quite different from other, well-established ERPs associated with prediction in language and can therefore nicely complement them. More specifically, they have at least two important differential features that make them interesting. The first is their latency, as they precede the predicted stimulus, rather than being triggered by it (as in the case of the N400 or the P600). The second is their morphology: they are relatively sustained and can build up several hundreds of milliseconds. Thus, future studies could capitalize on these characteristic features to investigate mechanisms of linguistic prediction at different time intervals (before and after the perception of pre-activated input) and with distinct time courses (shorter-lived and longer-lived cognitive processes).

Funding

This work was supported by The Swedish Research Council (Grant No.: 2018.00632 and 2021.00269), Knut and Alice Wallenberg Foundation (Grant No. 2018.0454), Crafoord Foundation (Grant No. 2017.0006), Marcus and Amalia Wallenberg Foundation (Grant No. 2018.0021). PLC was funded by Ministerio de Universidades of the Spanish Government (Margarita Salas para la formación de jóvenes doctores).

Declaration of Competing Interest

The authors declare no conflicts of interest, financial or otherwise.

Data availability

No data was used for the research described in the article.

Acknowledgements

We would like to thank Prof. Dr. Mireille Besson for her insightful feedback on an earlier version of the manuscript.

References

Allopenna, P.D., Magnuson, J.S., Tanenhaus, M.K., 1998. Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping models. J. Mem. Lang. *38* (4), 419–439. https://doi.org/10.1006/jmla.1997.2558.

Archibald, L.M.D., Joanisse, M.F., 2011. Electrophysiological responses to coarticulatory and word level miscues. J. Exp. Psychol.: Hum. Percept. Perform. *37* (4), 1275–1291. https://doi.org/10.1037/a0023506.

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

Beddor, P.S., McGowan, K.B., Boland, J.E., Coetzee, A.W., Brasher, A., 2013. The time course of perception of coarticulation. J. Acoust. Soc. Am. *133* (4), 2350–2366. https://doi.org/10.1121/1.4794366.

Bennett, D., Murawski, C., Bode, S., 2015. Single-trial event-related potential correlates of belief updating. ENEURO.0076-15.2015 eNeuro *2* (5). https://doi.org/10.1523/ENEURO.0076-15.2015.

Bentin, S., McCarthy, G., Wood, C.C., 1985. Event-related potentials, lexical decision and semantic priming. Electroencephalogr. Clin. Neurophysiol. *60* (4), 343–355. https://doi.org/10.1016/0013-4694(85)90008-2.

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. J. Exp. Psychol. Learn., Mem., Cogn. *31* (3), 443–467. https://doi.org/10.1037/0278-7393.31.3.443.

Besson, M., Faïta, F., 1995. An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. J. Exp. Psychol.: Hum. Percept. Perform. *21* (6), 1278–1296. https://doi.org/10.1037/0096-1523.21.6.1278.

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. Biol. Psychol. *46* (1), 3–23. https://doi.org/10.1016/S0301-0511(96)05215-5.

Bhangal, S., Sharma, S., Valle-Inclán, F., Ren, X., Hackley, S.A., 2021. Learning to deal with delayed outcomes: EEG oscillatory and slow potentials during the prefeedback interval. Psychophysiology *58* (9), e13853. https://doi.org/10.1111/psyp.13853.

Birbaumer, N., Elbert, T., Canavan, A.G., Rockstroh, B., 1990. Slow potentials of the cerebral cortex and behavior. Physiol. Rev. *70* (1), 1–41. 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. Int. J. Psychophysiol. *43* (1), 77–90. https://doi.org/10.1016/S0167-8760(01)00180-5.

Bornkessel-Schlesewsky, I., Schlesewsky, M., 2019. Toward a neurobiologically plausible model of language-related, negative event-related potentials. Front. Psychol. *10*, 298. https://doi.org/10.3389/fpsyg.2019.00298.

van Boxtel, G.J.M., Böcker, K.B.E., 2004. Cortical measures of anticipation. J. Psychophysiol. *18* (2/3), 61–76. https://doi.org/10.1027/0269-8803.18.23.61.

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., Wlotko, E.W., Warnke, L., Kuperberg, G.R., 2020. Going the extra mile: effects of discourse context on two late positivities during language comprehension. Neurobiol. Lang. *1* (1), 135–160. https://doi.org/10.1162/nol_a_00006.

Brown, C., Hagoort, P., 1993. The processing nature of the N400: evidence from masked priming. J. Cogn. Neurosci. *5* (1), 34–44. https://doi.org/10.1162/jocn.1993.5.1.34.

Bruce, G. (1977). Swedish word accents in sentence perspective. Gleerup.

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. Electroencephalogr. Clin. Neurophysiol. *69* (3), 234–243. https://doi.org/10.1016/0013-4694(88)90132-0.

Brunia, C.H.M., van Boxtel, G.J.M., Böcker, K.B.E., 2011. Negative slow waves as indices of anticipation: The bereitschaftspotential, the contingent negative variation, and the stimulus-preceding negativity. In: Kappenman, E.S., Luck, S.J. (Eds.), The Oxford Handbook of Event-Related Potential Components. Oxford University Press, p. 0. https://doi.org/10.1093/oxfordhb/9780195374148.013.0108.

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

Butler, S.R., Glass, A., Heffner, R., 1981. Asymmetries of the contingent negative variation (CNV) and its after positive wave (APW) related to differential hemispheric involvement in verbal and non-verbal task. Biol. Psychol. *13*, 157–171. https://doi.org/10.1016/0301-0511(81)90033-8.

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. Neurophysiol. Clin. /Clin. Neurophysiol. *44* (2), 189–201. https://doi.org/10.1016/j.neucli.2014.04.001.

Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. *36* (3), 181–204. https://doi.org/10.1017/S0140525×12000477.

Clausen, S.J., Kristensen, L.B., 2015. The cognitive status of stød. Nord. J. Linguist. *38* (2), 163–187. https://doi.org/10.1017/S0332586515000141.

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

Connolly, J.F., Phillips, N.A., 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. J. Cogn. Neurosci. *6* (3), 256–266. https://doi.org/10.1162/jocn.1994.6.3.256.

Coulson, S., King, J.W., Kutas, M., 1998. Expect the unexpected: event-related brain response to morphosyntactic violations. Lang. Cogn. Process. *13* (1), 21–58. https://doi.org/10.1080/016909698386582.

Dahan, D., Magnuson, J.S., Tanenhaus, M.K., Hogan, E.M., 2001. Subcategorical mismatches and the time course of lexical access: evidence for lexical competition. Lang. Cogn. Process. *16* (5–6), 507–534. https://doi.org/10.1080/01690960143000074.

Damen, E.J.P., Brunia, C.H.M., 1985. Slow brain potentials related to movement and visual feedback in a response timing task. Biol. Psychol. *20* (3), 195. https://doi.org/10.1016/0301-0511(85)90072-9.

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.

Dell, G.S., Chang, F., 2014. The P-chain: relating sentence production and its disorders to comprehension and acquisition. Philos. Trans. R. Soc. B: Biol. Sci. *369* (1634), 20120394. https://doi.org/10.1098/rstb.2012.0394.

DeLong, K.A., Urbach, T.P., Kutas, M., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Article 8. Nat. Neurosci. *8* (8) https://doi.org/10.1038/nn1504.

DeLong, K.A., Chan, W., Kutas, M., 2021. Testing limits: ERP evidence for word form preactivation during speeded sentence reading. Psychophysiology *58* (2), e13720. https://doi.org/10.1111/psyp.13720.

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

Dufour, S., Brunellière, A., Frauenfelder, U.H., 2013. Tracking the time course of word-frequency effects in auditory word recognition with event-related potentials. Cogn. Sci. *37* (3), 489–507. https://doi.org/10.1111/cogs.12015.

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. J. Mem. Lang. *41* (4), 469–495. https://doi.org/10.1006/jmla.1999.2660.

Federmeier, K.D., McLENNAN, D.B., Ochoa, E.D., 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., 2007. Multiple effects of sentential constraint on word processing. Brain Res. *1146*, 75–84. https://doi.org/10.1016/j.brainres.2006.06.101.

Fedorenko, E., Blank, I.A., 2020. Broca's area is not a natural kind. Trends Cogn. Sci. *24* (4), 270–284. https://doi.org/10.1016/j.tics.2020.01.001.

Ferreira, F., Chantavarin, S., 2018. Integration and prediction in language processing: a synthesis of old and new. Curr. Dir. Psychol. Sci. *27* (6), 443–448. https://doi.org/10.1177/0963721418794491.

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. J. Mem. Lang. *47* (2), 250–272. https://doi.org/10.1016/S0749-596X(02)00004-9.

Fischer-Jørgensen, E., 1989. Phonetic analysis of the stød in standard Danish. Phonetica *46* (1–3), 1–59. https://doi.org/10.1159/000261828.

Forster, K.I., 1981. Priming and the effects of sentence and lexical contexts on naming time: evidence for autonomous lexical processing. Q. J. Exp. Psychol. Sect. A *33* (4), 465–495. https://doi.org/10.1080/14640748108400804.

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 Lang. *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. Lang., Cogn. Neurosci. *31* (9), 1193–1205. https://doi.org/10.1080/23273798.2016.1205202.

Friederici, A.D., 2012. The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn. Sci. *16* (5), 262–268. https://doi.org/10.1016/j.tics.2012.04.001.

Frisch, S., Kotz, S.A., von Cramon, D.Y., Friederici, A.D., 2003. Why the P600 is not just a P300: the role of the basal ganglia. Clin. Neurophysiol. *114* (2), 336–340. https://doi.org/10.1016/S1388-2457(02)00366-8.

Friston, K., 2010. The free-energy principle: a unified brain theory? Article 2. Nat. Rev. Neurosci. *11* (2) https://doi.org/10.1038/nrn2787.

Friston, K.J., Rosch, R., Parr, T., Price, C., Bowman, H., 2018. Deep temporal models and active inference. Neurosci. Biobehav. Rev. *90*, 486–501. https://doi.org/10.1016/j.neubiorev.2018.04.004.

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.

Gómez, C.M., Arjona, A., Donnarumma, F., Maisto, D., Rodríguez-Martínez, E.I., Pezzulo, G., 2019. Tracking the time course of Bayesian inference with event-related potentials:a study using the central cue Posner paradigm. Front. Psychol. *10*, 1424. https://doi.org/10.3389/fpsyg.2019.01424.

Gosselke Berthelsen, S., Horne, M., Brännström, K.J., Shtyrov, Y., Roll, M., 2018. Neural processing of morphosyntactic tonal cues in second-language learners. J. Neurolinguist. *45*, 60–78. https://doi.org/10.1016/j.jneuroling.2017.09.001.

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*, e49562. 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. J. Neurosci. *37* (18), 4848–4858. https://doi.org/10.1523/JNEUROSCI.2800-16.2017.

Grisoni, L., Tomasello, R., Pulvermüller, F., 2021. Correlated brain indexes of semantic prediction and prediction error: Brain localization and category specificity. Cereb. Cortex *31* (3), 1553–1568. https://doi.org/10.1093/cercor/bhaa308.

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

Gunter, T.C., Friederici, A.D., Schriefers, H., 2000. Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. J. Cogn. Neurosci. *12* (4), 556–568. https://doi.org/10.1162/089892900562336.

Hagoort, P., 2017. The core and beyond in the language-ready brain. Neurosci. Biobehav. Rev. *81*, 194–204. https://doi.org/10.1016/j.neubiorev.2017.01.048.

Hagoort, P., Indefrey, P., 2014. The neurobiology of language beyond single words. Annu. Rev. Neurosci. *37* (1), 347–362. https://doi.org/10.1146/annurev-neuro-071013-013847.

Hagoort, P., Brown, C., Groothusen, J., 1993. The syntactic positive shift (sps) as an erp measure of syntactic processing. Lang. Cogn. Process. *8* (4), 439–483. https://doi.org/10.1080/01690969308407585.

Hagoort, P., Baggio, G., Willems, R.M., 2009. Semantic unification. In The cognitive neurosciences, 4th ed. MIT Press, pp. 819–836. ⟨https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_64579⟩.

Hale, J., 2016. Information-theoretical complexity metrics. Lang. Linguist. Compass *10* (9), 397–412. https://doi.org/10.1111/lnc3.12196.

Hanslmayr, S., Staudigl, T., Fellner, M.-C., 2012. Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. Front. Hum. Neurosci. *6*. ⟨https://www.frontiersin.org/articles/10.3389/fnhum.2012.00074⟩.

Heilbron, M., Armeni, K., Schoffelen, J.-M., Hagoort, P., de Lange, F.P., 2022. A hierarchy of linguistic predictions during natural language comprehension. Proc. Natl. Acad. Sci. *119* (32), e2201968119 https://doi.org/10.1073/pnas.2201968119.

Hjortdal, A., Frid, J., Roll, M., 2022. Phonetic and phonological cues to prediction: neurophysiology of Danish stød. J. Phon. *94*, 101178 https://doi.org/10.1016/j.wocn.2022.101178.

Hjortdal, A., Frid, J., Novén, M., Roll, M., 2024. Swift prosodic modulation of lexical access: brain potentials from three North Germanic language varieties. J. Speech, Lang., Hear. Res.: JSLHR 1–15. https://doi.org/10.1044/2023_JSLHR-23-00193.

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

Huang, Z., Feng, C., Qu, Q., 2023. Predicting coarse-grained semantic features in language comprehension: evidence from ERP representational similarity analysis and Chinese classifier. Cereb. Cortex *33* (13), 8312–8320. https://doi.org/10.1093/cercor/bhad116.

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

Huettig, F., Guerra, E., 2019. Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. Brain Res. *1706*, 196–208. https://doi.org/10.1016/j.brainres.2018.11.013.

Huettig, F., Mani, N., 2016. Is prediction necessary to understand language? Probably not. Lang. Cogn. Neurosci. *31* (1), 19–31. https://doi.org/10.1080/23273798.2015.1072223.

Huettig, F., Audring, J., Jackendoff, R., 2022. A parallel architecture perspective on pre-activation and prediction in language processing. Cognition *224*, 105050. https://doi.org/10.1016/j.cognition.2022.105050.

Hunter, C.R., 2013. Early effects of neighborhood density and phonotactic probability of spoken words on event-related potentials. Brain Lang. *127* (3), 463–474. https://doi.org/10.1016/j.bandl.2013.09.006.

Ito, A., Corley, M., Pickering, M.J., Martin, A.E., Nieuwland, M.S., 2016. Predicting form and meaning: evidence from brain potentials. J. Mem. Lang. *86*, 157–171. https://doi.org/10.1016/j.jml.2015.10.007.

Jakuszeit, M., Kotz, S.A., Hasting, A.S., 2013. Generating predictions: lesion evidence on the role of left inferior frontal cortex in rapid syntactic analysis. Cortex *49* (10), 2861–2874. https://doi.org/10.1016/j.cortex.2013.05.014.

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

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

Koester, D., Gunter, Th.C., Wagner, S., Friederici, A.D., 2004. Morphosyntax, prosody, and linking elements: the auditory processing of German nominal compounds. J. Cogn. Neurosci. *16* (9), 1647–1668. https://doi.org/10.1162/0898929042568541.

Kornhuber, H.H., Deecke, L., 1965. Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: bereitschaftspotential und reafferente Potentiale. PflüGer. 'S. Arch. F. üR. die Gesamt Physiol. Des. Mensch 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. Biol. Psychol. *50* (1), 1–18. https://doi.org/10.1016/S0301-0511(98)00047-7.

Kuperberg, G.R., 2013. The proactive comprehender: What event-related potentials tell us about the dynamics of reading comprehension. In: Miller, B., Cuttling, L.E., McCardle, P. (Eds.), Unraveling the Behavioral, Neurobiological, and Genetic Components of Reading Comprehension. Paul Brookes Publishing, pp. 176–192.

Kuperberg, G.R., 2021. Tea With milk? A hierarchical generative framework of sequential event comprehension. Top. Cogn. Sci. *13* (1), 256–298. https://doi.org/10.1111/tops.12518.

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

Kuperberg, G.R., Sitnikova, T., Caplan, D., Holcomb, P.J., 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. Cogn. Brain Res. *17* (1), 117–129. https://doi.org/10.1016/S0926-6410(03)00086-7.

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. J. Cogn. Neurosci. *32* (1), 12–35. https://doi.org/10.1162/jocn_a_01465.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. *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. Article 5947. Nature *307* (5947) https://doi.org/10.1038/307161a0.

Kutas, M., King, J.W., 1996. In: Toshio, I., McClelland, J.L. (Eds.), The potentials for basic science processing: Differentiating integrative processes. The MIT Press, pp. 501–546. https://doi.org/10.7551/mitpress/1479.003.0031.

Kutas, M., van Petten, C., Kluender, R., 2006. Chapter 17. Psycholinguistics Electrified II (1994–2005). Handb. Psycholinguist. 659–724. https://doi.org/10.1016/B978-012369374-7/50018-3.

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

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

Lau, E., Stroud, C., Plesch, S., Phillips, C., 2006. The role of structural prediction in rapid syntactic analysis. Brain Lang. *98* (1), 74–88. https://doi.org/10.1016/j.bandl.2006.02.003.

Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (De) constructing the N400. Article 12. Nat. Rev. Neurosci. *9* (12) https://doi.org/10.1038/nrn2532.

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), 13. https://doi.org/10.1525/collabra.40.

Lehmann, D., Skrandies, W., 1980. Visually evoked scalp potential fields in hemiretinal stimulation. In: Schmöger, E., Kelsey, J.H. (Eds.), Visual Electrodiagnosis in Systemic Diseases: Proceedings of the 17th I.S.C.E.V. Symposium Erfurt, June 5–10, 1979. Springer, Netherlands, pp. 237–243. https://doi.org/10.1007/978-94-009-9180-4_35.

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.

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. https://doi.org/10.1016/j.neuroimage.2019.01.005.

León-Cabrera, P., Pagonabarraga, J., Morís, J., Martínez-Horta, S., Marín-Lahoz, J., Horta-Barba, A., Bejr-Kasem, H., Kulisevsky, J., Rodríguez-Fornells, A., 2021. Neural signatures of predictive language processing in Parkinson's disease with and without mild cognitive impairment. Cortex *141*, 112–127. https://doi.org/10.1016/j.cortex.2021.03.032.

León-Cabrera, P., Piai, V., Morís, J., Rodríguez-Fornells, A., 2022. Alpha power decreases associated with prediction in written and spoken sentence comprehension. Neuropsychologia *173*, 108286. https://doi.org/10.1016/j.neuropsychologia.2022.108286.

Levy, R., 2008. A noisy-channel model of human sentence comprehension under uncertain input. In: Lapata, M., Ng, H.T. (Eds.), Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, pp. 234–243. ⟨https://aclanthology.org/D08-1025⟩.

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. https://doi.org/10.1016/j.neuropsychologia.2017.05.017.

Lozano-Argüelles, C., Sagarra, N., Casillas, J.V., 2020. Slowly but surely: interpreting facilitates L2 morphological anticipation based on suprasegmental and segmental information. Biling.: Lang. Cogn. *23* (4), 752–762. https://doi.org/10.1017/S1366728919000634.

Lupyan, G., Clark, A., 2015. Words and the world: predictive coding and the language-perception-cognition interface. Curr. Dir. Psychol. Sci. *24* (4), 279–284. https://doi.org/10.1177/0963721415570732.

Macar, F., Vidal, F., 2004. Event-related potentials as indices of time processing: a review. J. Psychophysiol. 18 (2/3), 89–104. https://doi.org/10.1027/0269-8803.18.23.89.

Marslen-Wilson, W.D., 1987. Functional parallelism in spoken word-recognition. Cognition *25* (1), 71–102. https://doi.org/10.1016/0010-0277(87)90005-9.

Matchin, W., Hammerly, C., Lau, E., 2017. The role of the IFG and pSTS in syntactic prediction: evidence from a parametric study of hierarchical structure in fMRI. Cortex *88*, 106–123. https://doi.org/10.1016/j.cortex.2016.12.010.

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

McClelland, J.L., 2013. Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. J. Exp. Psychol. Gen. *142* (4), 1190–1210. https://doi.org/10.1037/a0033812.

McClelland, J.L., Elman, J.L., 1986. The TRACE model of speech perception. Cogn. Psychol. *18* (1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0.

Michaelov, J.A., Bardolph, M.D., Coulson, S., Bergen, B.K., 2021. Different kinds of cognitive plausibility: Why are transformers better than RNNs at predicting N400 amplitude? (arXiv:2107.09648). arXiv. https://doi.org/10.48550/arXiv.2107.09648.

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.J., 2008. An electrophysiological investigation of early effects of masked morphological priming. Lang. Cogn. Process. *23* (7–8), 1021–1056. https://doi.org/10.1080/01690960802299386.

Morton, J., 1969. Interaction of information in word recognition. Psychol. Rev. *76* (2), 165–178. https://doi.org/10.1037/h0027366.

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.

Neville, H., Nicol, J.L., Barss, A., Forster, K.I., Garrett, M.F., 1991. Syntactically based sentence processing classes: evidence from event-related brain potentials. J. Cogn. Neurosci. *3* (2), 151–165. https://doi.org/10.1162/jocn.1991.3.2.151.

Nieuwland, M.S., Barr, D.J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D.I., Ferguson, H.J., Fu, X., Heyselaar, E., Huettig, F., Matthew Husband, E., Ito, A., Kazanina, N., Kogan, V., Kohút, Z., Kulakova, E., Mézière, D., Politzer-Ahles, S., Rousselet, G., Von Grebmer Zu Wolfsthurn, S., 2020. Dissociable effects of prediction and integration during language comprehension: evidence from a large-scale study using brain potentials. Philos. Trans. R. Soc. Lond. Ser. B, Biol. Sci. *375* (1791), 20180522. https://doi.org/10.1098/rstb.2018.0522.

Norris, D., 1986. Word recognition: context effects without priming. Cognition *22* (2), 93–136. https://doi.org/10.1016/s0010-0277(86)90001-6.

Norris, D., McQueen, J.M., Cutler, A., 2016. Prediction, Bayesian inference and feedback in speech recognition. Lang. Cogn. Neurosci. *31* (1), 4–18. https://doi.org/10.1080/23273798.2015.1081703.

Novén, M. (2021). Brain anatomical correlates of perceptual phonological proficiency and language learning aptitude [Thesis/doccomp, Lund University]. ⟨http://lup.lub.lu.se/record/69158b8b-8a66-4ebf-9327-b59d0f9422d1⟩.

Novick, J.M., Trueswell, J.C., Thompson-Schill, S.L., 2010. Broca's area and language processing: evidence for the cognitive control connection. Lang. Linguist. Compass *4* (10), 906–924. https://doi.org/10.1111/j.1749-818X.2010.00244.x.

Osterhout, L., 1999. A superficial resemblance does not necessarily mean you are part of the family: counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. Lang. Cogn. Process. *14* (1), 1–14. https://doi.org/10.1080/016909699386356.

Osterhout, L., Holcomb, P.J., 1992. Event-related brain potentials elicited by syntactic anomaly. J. Mem. Lang. *31* (6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z.

Patel, A.D., Gibson, E., Ratner, J., Besson, M., Holcomb, P.J., 1998. Processing syntactic relations in language and music: an event-related potential study. J. Cogn. Neurosci. *10* (6), 717–733. https://doi.org/10.1162/089892998563121.

van Petten, C., 1995. Words and sentences: event-related brain potential measures. Psychophysiology *32* (6), 511–525. https://doi.org/10.1111/j.1469-8986.1995.tb01228.x.

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

van Petten, C., Coulson, S., Rubin, S., Plante, E., Parks, M., 1999. Time course of word identification and semantic integration in spoken language. J. Exp. Psychol. Learn. Mem. Cogn. *25* (2), 394–417. https://doi.org/10.1037//0278-7393.25.2.394.

Pickering, M.J., Garrod, S., 2013. An integrated theory of language production and comprehension. Behav. Brain Sci. *36* (4), 329–347. https://doi.org/10.1017/S0140525×12001495.

Poli, S., Sarlo, M., Bortoletto, M., Buodo, G., Palomba, D., 2007. Stimulus-preceding negativity and heart rate changes in anticipation of affective pictures. Int. J. Psychophysiol. *65* (1), 32–39. https://doi.org/10.1016/j.ijpsycho.2007.02.008.

Proverbio, A.M., Crotti, N., Zani, A., Adorni, R., 2009. The role of left and right hemispheres in the comprehension of idiomatic language: an electrical neuroimaging study. BMC Neurosci. *10* (1), 116. https://doi.org/10.1186/1471-2202-10-116.

Pulvermüller, F., Grisoni, L., 2020. Semantic prediction in brain and mind. Trends Cogn. Sci. *24* (10), 781–784. https://doi.org/10.1016/j.tics.2020.07.002.

Rabovsky, M., Hansen, S.S., McClelland, J.L., 2018. Modelling the N400 brain potential as change in a probabilistic representation of meaning. Article 9. Nat. Hum. Behav. *2* (9) 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. J. Exp. Psychol. Hum. Percept. Perform. *37* (2), 514–528. https://doi.org/10.1037/a0020990.

Rebert, C.S., Lowe, R.C., 1980. Task-related hemispheric asymmetry of contingent negative variation (Scopus). Prog. Brain Res. 54 (C), 776–781. https://doi.org/10.1016/S0079-6123(08)61702-1.

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.

Roll, M., 2015. A neurolinguistic study of South Swedish word accents: Electrical brain potentials in nouns and verbs. Nord. J. Linguist. *38* (2), 149–162. https://doi.org/10.1017/S0332586515000189.

Roll, M., 2022. The predictive function of Swedish word accents. Front. Psychol. *13*. 〈https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.910787〉.

Roll, M., Horne, M., Lindgren, M., 2010. Word accents and morphology—ERPs of Swedish word processing. Brain Res. *1330*, 114–123. https://doi.org/10.1016/j.brainres.2010.03.020.

Roll, M., Söderström, P., Horne, M., 2013. Word-stem tones cue suffixes in the brain. Brain Res. *1520*, 116–120. https://doi.org/10.1016/j.brainres.2013.05.013.

Roll, M., Söderström, P., Mannfolk, P., Shtyrov, Y., Johansson, M., van Westen, D., Horne, M., 2015. Word tones cueing morphosyntactic structure: neuroanatomical substrates and activation time-course assessed by EEG and fMRI. Brain Lang. *150*, 14–21. https://doi.org/10.1016/j.bandl.2015.07.009.

Roll, M., Söderström, P., Frid, J., Mannfolk, P., Horne, M., 2017. Forehearing words: pre-activation of word endings at word onset. Neurosci. Lett. *658*, 57–61. https://doi.org/10.1016/j.neulet.2017.08.030.

Roll, M., Söderström, P., Horne, M., Hjortdal, A., 2023. Pre-activation negativity (PrAN): a neural index of predictive strength of phonological cues. Article 1. Lab. Phonol. *14* (1) https://doi.org/10.16995/labphon.6438.

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

Rösler, F., Heil, M., Röder, B., 1997. Slow negative brain potentials as reflections of specific modular resources of cognition. Biol. Psychol. *45* (1), 109–141. https://doi.org/10.1016/S0301-0511(96)05225-8.

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

Rugg, M.D., 1984b. Event-related potentials in phonological matching tasks. Brain Lang. *23* (2), 225–240. https://doi.org/10.1016/0093-934X(84)90065-8.

Rugg, M.D., 1984a. Event-related potentials and the phonological processing of words and non-words. Neuropsychologia *22* (4), 435–443. https://doi.org/10.1016/0028-3932(84)90038-1.

Sagarra, N., Casillas, J.V., 2018. Suprasegmental information cues morphological anticipation during L1/L2 lexical access. J. Second Lang. Stud. *1* (1), 31–59. https://doi.org/10.1075/jsls.17026.sag.

Salverda, A.P., Kleinschmidt, D., Tanenhaus, M.K., 2014. Immediate effects of anticipatory coarticulation in spoken-word recognition. J. Mem. Lang. *71* (1), 145–163. https://doi.org/10.1016/j.jml.2013.11.002.

Schwanenflugel, P.J., LaCount, K.L., 1988. Semantic relatedness and the scope of facilitation for upcoming words in sentences. J. Exp. Psychol.: Learn., Mem., Cogn. *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. J. Mem. Lang. *24* (2), 232–252. https://doi.org/10.1016/0749-596X(85)90026-9.

Silcox, J.W., Mickey, B., Payne, B.R., 2023. Disruption to left inferior frontal cortex modulates semantic prediction effects in reading and subsequent memory: evidence from simultaneous TMS-EEG. Psychophysiology *60* (9), e14312. https://doi.org/10.1111/psyp.14312.

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

Söderström, P., Roll, M., Horne, M., 2012. Processing morphologically conditioned word accents. Ment. Lex. *7* (1), 77–89. https://doi.org/10.1075/ml.7.1.04soe.

Söderström, P., Horne, M., Frid, J., Roll, M., 2016. Pre-activation negativity (PrAN) in brain potentials to unfolding words. Front. Hum. Neurosci. *10*. 〈https://www.frontiersin.org/articles/10.3389/fnhum.2016.00512〉.

Söderström, P., Horne, M., Roll, M., 2017. Stem tones pre-activate suffixes in the brain. J. Psycholinguist. Res. *46* (2), 271–280. https://doi.org/10.1007/s10936-016-9434-2.

Söderström, P., Horne, M., Mannfolk, P., van Westen, D., Roll, M., 2017. Tone-grammar association within words: concurrent ERP and fMRI show rapid neural pre-activation and involvement of left inferior frontal gyrus in pseudoword processing. Brain Lang. *174*, 119–126. https://doi.org/10.1016/j.bandl.2017.08.004.

Söderström, P., Horne, M., Mannfolk, P., van Westen, D., Roll, M., 2018. Rapid syntactic pre-activation in Broca's area: concurrent electrophysiological and haemodynamic recordings. Brain Res. *1697*, 76–82. https://doi.org/10.1016/j.brainres.2018.06.004.

Squires, N.K., Squires, K.C., Hillyard, S.A., 1975. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. Electroencephalogr. Clin. Neurophysiol. *38* (4), 387–401. https://doi.org/10.1016/0013-4694(75)90263-1.

Staub, A., 2015. The effect of lexical predictability on eye movements in reading: critical review and theoretical interpretation. Lang. Linguist. Compass *9* (8), 311–327. https://doi.org/10.1111/lnc3.12151.

Taylor, W.L., 1953. Cloze procedure": a new tool for measuringreadability. J. Q. *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. Article 1. Sci. Rep. *9* (1) https://doi.org/10.1038/s41598-019-56600-x.

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. Proc. Natl. Acad. Sci. *94* (26), 14792–14797. https://doi.org/10.1073/pnas.94.26.14792.

Traxler, M.J., Foss, D.J., 2000. Effects of sentence constraint on priming in natural language comprehension. J. Exp. Psychol. Learn., Mem., Cogn. *26* (5), 1266–1282. https://doi.org/10.1037//0278-7393.26.5.1266.

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

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

Walter, W.G., Cooper, R., Aldridge, V.J., Mccallum, W.C., Winter, A.L., 1964. Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. Nature *203*, 380–384. https://doi.org/10.1038/203380a0.

Wang, L., Kuperberg, G., Jensen, O., 2018. Specific lexico-semantic predictions are associated with unique spatial and temporal patterns of neural activity. eLife *7*, e39061. https://doi.org/10.7554/eLife.39061.

Wang, L., Hagoort, P., Jensen, O., 2018. Language prediction Is reflected by coupling between frontal gamma and posterior alpha oscillations. J. Cogn. Neurosci. *30* (3), 432–447. https://doi.org/10.1162/jocn_a_01190.

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. J. Cogn. Neurosci. *16* (7), 1272–1288. https://doi.org/10.1162/0898929041920487.

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.

Zhuang, J., Tyler, L.K., Randall, B., Stamatakis, E.A., Marslen-Wilson, W.D., 2014. Optimally efficient neural systems for processing spoken language. Cereb. Cortex *24* (4), 908–918. https://doi.org/10.1093/cercor/bhs366.