



LUND UNIVERSITY

Pre-activation negativity (PrAN)

A neural index of predictive strength of phonological cues

Roll, Mikael; Söderström, Pelle; Hjortdal, Anna; Horne, Merle

Published in:
Laboratory Phonology

2023

Document Version:
Peer reviewed version (aka post-print)

[Link to publication](#)

Citation for published version (APA):

Roll, M., Söderström, P., Hjortdal, A., & Horne, M. (Accepted/In press). Pre-activation negativity (PrAN): A neural index of predictive strength of phonological cues. *Laboratory Phonology*, 1-28.

Total number of authors:
4

General rights

Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Pre-activation negativity (PrAN): A neural index of predictive strength of phonological cues

Mikael Roll¹, Pelle Söderström^{1,2}, Anna Hjortdal¹, and Merle Horne¹

¹Centre for Languages and Literature, Lund University, Sweden

²MARCS Institute for Brain, Behaviour & Development, Western Sydney University, Australia

Corresponding author: Mikael Roll (mikael.roll@ling.lu.se)

Abstract

We propose that a recently discovered event-related potential (ERP) component—the pre-activation negativity (PrAN)—indexes the predictive strength of phonological cues, including segments, word tones, and sentence-level tones. Specifically, we argue that PrAN is a reflection of the brain’s anticipation of upcoming speech (segments, morphemes, words, and syntactic structures). Findings from a long series of neurolinguistic studies indicate that the effect can be divided into two time windows with different possible brain sources. Between 136 and 200 ms from stimulus onset, it indexes activity mainly in the primary and secondary auditory cortices, reflecting disinhibition of neurons sensitive to the expected acoustic signal, as indicated by the brain regions’ response to predictive certainty rather than sound salience. After ~200 ms, PrAN is related to activity in Broca’s area, possibly reflecting inhibition of irrelevant segments, morphemes, words, and syntactic structures.

Keywords: prediction, phonology, prosody, brain, event-related potentials

1. Introduction

This article reviews recent neurophysiological results providing evidence that phonological cues, both segmental and prosodic, give rise to a negativity in event-related potentials (ERPs) which increases in amplitude as a function of the predictive strength of the cue with respect to upcoming linguistic information. Such cues are expected given the fact that the physical, social, and communicative environment in which humans develop and interact does not change randomly but, rather, predictably. The *predictive coding* framework posits that the brain processes sensory information by maintaining a model of the world that generates hypotheses about the immediate future, enabling us to interact rapidly with the environment and minimize energy expenditure (Friston, 2005; Heilbron & Chait, 2018; Rao & Ballard, 1999). The perception system helps refine the model through the active sampling of the environment by pre-activating expected sensory information and continuously reporting the *prediction error*, that is, to what extent perception does *not* conform to the anticipations (Friston, 2018; Friston & Kiebel, 2009). Within the hierarchical predictive coding framework, predictions are propagated downwards from cognitively higher to lower cortical areas. Only input differing from predictions is passed forward as prediction error (Rao & Ballard, 1999).

There is no a priori reason to believe that the brain treats language any differently from other information (Friston et al., 2021; Yildiz et al., 2013). In line with this, evidence is accumulating about the role of prediction in language processing (Gagnepain et al., 2012; Kuperberg & Jaeger, 2016). Studies on event-related potentials (ERP) have mainly explored what can be argued to be the effects of *prediction error* (encountering unexpected stimuli) and updating the internal model of the world (*belief updating*) (DeLong et al., 2005; Federmeier, 2007). However, the actual *pre-activation* of linguistic representations before they are perceived would also likely leave measurable traces in the ERP signal. This pre-activation would, to a large extent, be based on phonological cues and can be thought to vary with the predictive strength of these cues. Recently, the *pre-activation negativity* (PrAN) has been proposed to reflect the phonological cues' pre-activation of word forms (Roll et al., 2017; Roll et al., 2015; Söderström, Horne, Frid, et al., 2016) and syntactic structures (Söderström et al., 2018). The present article suggests that PrAN reflects the predictive strength of phonological cues. It also discusses possible neural sources of prediction-related potentials.

The article is organized in the following way. Section 2 will discuss the language-internal properties that make predictive processing effective. After that, Section 3 will briefly review the most common way of measuring the neurophysiological correlates of predictive strength: Neural responses to failed predictions. Section 4 will show the neurophysiological indexes of the actual prediction as it is being formed. In this context, the pre-activation negativity will be presented. The subsequent sections will describe PrAN in response to suprasegmental cues (Section 5) and discuss its relation to other ERP components (Section 6). Sections 7 and 8 will touch upon the possible neural sources of PrAN and shared neural traits of prediction and prediction error. Finally, Section 9 will wrap up the information from the other sections and draw some general conclusions.

2. Language-internal pressure on prediction

The main goal of the brain has been argued to be minimizing prediction error, that is, avoiding surprise in both the short and long run (Friston & Kiebel, 2009). Therefore, external cues for anticipation that are more reliable and thus lead to increased certainty also elevate pre-activation at lower levels of neural representation (Friston, 2005). In information theory, the uncertainty about the immediate future at a certain stage is measurable as *entropy* (Shannon, 1948). Entropy can also be understood as the expected *surprisal* or prediction error of the outcome of an event (Gwilliams & Davis, 2022). We can therefore say that the lower the entropy (uncertainty) at a specific point in the processing of a spoken word, the more listeners can commit to the possible continuations of a word beginning (Ettinger et al., 2014) and begin to pre-activate those continuations. Entropy can be calculated from the number of possible outcomes of an event and their respective probability. In the case of a word beginning, the outcomes are the *lexical competitors* conforming possible word completions, and the probability can be approximated by each competitor's relative frequency of occurrence. Both factors have been extensively covered in the psycholinguistic spoken word-recognition literature (Marslen-Wilson & Tyler, 1980; McClelland & Elman, 1986; Norris, 1994; Norris & McQueen, 2008; Norris et al., 2016). If there are fewer possible word candidates, we can be more confident that one of them will occur. Hence, those particular words will be strongly activated. In spoken word recognition, a number of possible outcomes at the beginning of a word compete for activation in a process referred to as *lexical competition*. In English, the phoneme sequence /dʒɛ/—as in *gender* or *jetlag*—activates a larger number of possible word competitors compared to /zɛ/ in *zealot*. The *smaller* the lexical competition of a sequence of word-initial phonemes during online listening, the more certain a listener can be about how the word is going to end. The most relevant lexical competitors are based on the mental dictionary of the listener. However, since experimenters do not have direct access to that dictionary, lexical competition can be calculated using pronunciation-based lexicons and

corpora. In the above example, the phoneme sequence / $\widehat{d}z\epsilon$ -/ has almost 11 times as many possible continuations as / $z\epsilon$ -/ in the English Lexicon Project corpus (Balota et al., 2007).

In our studies on Swedish, we used the NST lexicon (Andersen, 2011) and PAROLE¹ to calculate lexical statistics. For instance, if—in a context where a noun is expected—a native Swedish listener hears a word beginning with the phoneme sequence / fy /, whose lexical competition consists of 52 possible continuations, (s)he can be relatively sure that one of the few possibilities (for example *fyndet* ‘the finding’) will follow. If, instead, the listener hears the phoneme sequence / fa /, which has 657 lexical competitors (as in *fallet* ‘the case’), the certainty is much lower due to the high number of possibilities. The frequency of occurrence of the competitors should also influence prediction. It can be approximated by the word frequency in the corpus. It seems intuitive that—in the absence of context—a listener would more strongly expect a frequent word than an infrequent one. Indeed, mathematically, increased word frequency of some of a word beginning’s lexical competitors also lowers the entropy. Therefore, word beginnings involving some more frequent competitors should create stronger pre-activation than word beginnings cueing less frequent lexical competitors. This factor might well work against the lexical competition effect so that word-initial phonemes evoking a relatively large number of competitors—which would normally give weak pre-activation—might yield stronger pre-activation if some of the competitors are high-frequency words. That would indeed be the case in the example above: whereas the 52 competitors beginning with / fy / occur on average 16 times each, the 657 competitors starting on / fa / are much more common, occurring on average 37 times each in the corpus. The difference in frequency would be expected to slightly adjust the predictive strength in favor of the otherwise weaker predictor / fa /.

Attention is another factor that is involved in predictive processing and is triggered partly by language-internal properties. Informally, paying more attention to a certain part of the speech signal will give rise to a stronger neural reaction. In the predictive coding framework, attention is formalized as the gain level of prediction error units. The gain is modulated by the anticipated *precision* of predictions, that is, the confidence we place in our expectations (Friston, 2009). Specifically, when we expect our generated predictions to be more precise, prediction error becomes more informative. We can then increase the prediction error gain to allow the prediction error to influence the predictive model more and make future generated predictions more precise. Auditory attention has spectral, temporal, and spatial dimensions. At specific points in time and space, we expect spectral predictions to be more precise and, therefore, increase the prediction error gain/attention (Nobre & van Ede, 2018). For example, by getting attuned to a rhythmic pattern from a particular sound source, we can have more exact spectral expectations at the beats than between them, leading to increased prediction error gain at the beats (cf. Fitzroy & Sanders, 2015, 2021). In a stress-timed language such as Swedish or English, this means increased attention/prediction error gain in stressed syllables, often coinciding with word onsets (Astheimer & Sanders, 2009, 2011) due to the predominant trochaic rhythm. The enhanced prediction error gain at word onset can further aid speech segmentation and lexical access (Cutler & Norris, 1988), as developed in the next section.

3. Neurophysiological effects of prediction error

There is rich evidence from language studies for what can be argued to be online neural measures of prediction error or surprise at detecting an unexpected stimulus and the subsequent update of the internal model of the environment, also referred to as *belief updating*. Auditory stimuli produce a large characteristic pattern in the ERPs. A major negative peak between 80–110 ms is called *N1*, and the following positive peak, *P2* (160–200 ms) (Davis,

¹ <http://spraakdata.gu.se/parole/lexikon/swedish.parole.lexikon.html>

1964; Hillyard et al., 1973). The N1 is found for word onsets in isolated words and connected speech (Sanders & Neville, 2003), and has its sources in auditory cortex (Näätänen & Picton, 1987). Since N1 amplitude is larger for unpredicted stimuli, it has been suggested to relate to prediction error (Schröger et al., 2015). When listeners control the occurrence of stimuli themselves, or if the stimulus is self-generated, the component is substantially reduced (N1 suppression) (McCarthy & Donchin, 1976; Schafer & Marcus, 1973), due to minimized prediction error (Hsu et al., 2016). The N1 is also regulated by attention (Hillyard et al., 1973). In the time dimension, this means that the peak increases for stimuli delivered at attended moments (Lange & Röder, 2006; Lange et al., 2003; Sanders & Astheimer, 2008). Importantly for speech processing, rhythmically strong positions increase the N1 independently of loudness (Fitzroy & Sanders, 2015, 2021). This kind of temporal attention likely underlies N1 enhancement for both linguistic and non-linguistic stimuli at word onsets (Astheimer & Sanders, 2009). Astheimer and Sanders (2011) argued that attention is directed to word onsets because the information is unpredictable, and therefore, more resources are needed to process them. They manipulated the predictability of word onsets in an artificial language-learning study by letting some words always appear after others, making them highly predictable. The predictable word onsets did not increase the N1 after training, whereas the unpredictable word onsets did.

A reinterpretation of Astheimer and Sanders's (2011) proposal in terms of predictive coding could be that two factors enhance the N1 at word-initial and stressed syllables: the amount and the gain of prediction error. The first factor concerns the uncertainty about the form and content of stressed syllables. Specifically, prediction error would typically be greater upon hearing stressed syllables because the entropy – or expected surprisal – is higher. Rhythmically strong syllables in English are generally more informative than weak ones because they involve more options. They often correspond to a word onset (Cutler & Norris, 1988) or root syllable. This is where the lexical open-class information is found, and the number of possible continuations is extensive before the syllable starts unfolding. The prediction error at these points is useful for minimizing future surprise: The lexically strong points help the listener refine coming predictions while building up a semantic context. Therefore, it is advantageous for the system to increase the prediction error gain at the strong points. Weak syllables, conversely, tend to represent closed-class categories, such as grammatical endings or function words. Hence, the entropy is higher at time points where stressed syllables are predicted to occur, increasing the average prediction error upon actually hearing them and, therefore, the N1.

The second factor has to do with attention, which Friston (2009) models as prediction error gain due to the predicted *specificity* of a stimulus. Concerning speech, this can be understood as follows. Stressed syllables need to be phonologically more *specific* to separate the vast number of options. They can contain any of 19 contrasting vowel sounds in Standard Southern British English. Weak syllables, in contrast, are reduced to the two centralized vowels [ə] or [ɪ]. The predicted increase in specificity at a stressed syllable invites the listener to raise the prediction error gain. Even smaller acoustic deviances from what is expected will carry important information and should be allowed to influence the predictive model. This view of specificity is supported by the fact that phonetic reduction is proportional to the probability of words (Cohen, 2014; Jurafsky et al., 2001). Production and perception seem to go hand in hand. The speaker increases the specificity at informative points in the speech. The listener follows by predicting specificity to be higher at these points and raising the prediction error gain or, in other words, the attention. “N1 suppression rebounds” in predictive sound sequences might also be explained in terms of the rising specificity of predictions leading to accumulated prediction error gain (Hsu et al., 2016). Prosodic cues that increase the negativity in the N1 time range add to the interpretation of the N1 in terms of prediction error since they are unexpected in their context (Mietz et al., 2008; Roll & Horne, 2011; Roll et al., 2013). Brain activity in the P2-component time range has been argued to

reflect *passive anticipatory attention* (Roll & Horne, 2011; Roll et al., 2013). We will slightly reinterpret this below as meaning increased predictive allocation of resources to a receptive field in the auditory cortex (cf. Nobre & van Ede, 2018). However, we will argue that it indexes not necessarily a gain increase (attention) but rather a reweighting of the predictive model by disinhibition of the neurons anticipated to be relevant for processing future incoming auditory features (Almeida, 2021; Garrett et al., 2020).

The mismatch negativity (MMN) family of neural responses has also been proposed to constitute measures of prediction error (Friston, 2005; Wacongne et al., 2012; Wacongne et al., 2011) or belief update (Friston et al., 2021). The main differences between N1 and MMN are their timing—MMN usually has a later time window—and the fact that MMN is normally only reported in relation to the ‘oddball paradigm.’ The MMN is a negative ERP component typically occurring between 100 to 250 ms following stimuli that are unexpected due to their low frequency of occurrence within an oddball paradigm experiment, where repeated presentations of a frequent (*standard*) stimulus are interspersed with the occasional delivery of an infrequent (*deviant*) stimulus. The brain will expect the standard to a higher degree than the deviant stimulus. This leads to prediction error when the deviant is presented. The prediction error gives rise to an MMN, originating in the “prediction error layer”, layer 4, of the auditory cortex (Wacongne et al., 2012).

A negative deflection has been found following unexpected speech sounds without using an oddball paradigm. It is referred to as the *phonological mapping negativity* (PMN, originally phonological *mismatch negativity*) and has been observed in paradigms where an expectation for a certain word form has been created. A PMN is generated if a stimulus word onset does not acoustically match the anticipated word form (Connolly & Phillips, 1994). The PMN has been reported consistently at centroanterior electrodes in time windows between ~220 and 350 ms (Connolly & Phillips, 1994; Connolly et al., 1992; Connolly et al., 2001; Connolly et al., 1990; Newman & Connolly, 2009; Newman et al., 2003; van den Brink et al., 2001). Before 200 ms, a centroposterior effect has also been described with onsets at 130–140 ms (D’Arcy et al., 2000; van den Brink et al., 2001). Interestingly, the studies finding PMN effects occurring mainly before 200 ms have in common that the onset consonant of the PMN-eliciting words drastically changed an acoustically specific expectation built up under relatively naturalistic conditions. D’Arcy et al. (2000) elicited an early PMN by mismatching limited options in the description of a previously presented visual scene. In a similar way, van den Brink et al. (2001) used sentences where a word was strongly expected due to high cloze probability. An early PMN was produced by words where the onset phonemes mismatched the predicted word. Studies with later PMN increase have investigated word onsets that occurred in semantically less constraining contexts (Connolly et al., 1992; Connolly et al., 1990), onsets that did not match the highest cloze probability word (Connolly & Phillips, 1994), or unfulfilled expectations formed by instructing participants to alter the onset consonant of a stimulus word (Connolly et al., 2001; Kujala et al., 2004; Newman & Connolly, 2009; Newman et al., 2003). The later PMN has been source-localized to the left frontal lobe using ERPs (Connolly et al., 2001) and to the left anterior temporal lobe using magnetoencephalography (MEG) (Kujala et al., 2004). A study with a PMN onset slightly before 200 ms presented a picture and then a word describing it or not. A frontal negativity was found at 180–280 ms (Duta et al., 2012). The PMN has been observed for unpredicted speech sounds independent of their lexicality (Newman & Connolly, 2009).

Another early negative component, the early left-anterior negativity (ELAN)—responding to unexpected morphological or syntactic structures—has previously been related to the MMN (Pulvermüller & Shtyrov, 2003), and might receive a similar explanation in terms of prediction error. A negative peak occurring in a later time window that has been linked to prediction error at a higher cognitive level is the N400 (Almeida, 2021; Bornkessel-Schlesewsky & Schlesewsky, 2019). Updating the current predictive model (*belief updating*) has been chiefly associated with somewhat later, positive deflections related to the P3

(Donchin & Coles, 1988; Friston et al., 2021) and P600 components (Sassenhagen et al., 2014). The P600 is traditionally said to reflect syntactic and morphological reanalysis (Osterhout & Holcomb, 1992; Rodriguez-Fornells et al., 2001).

4. Neurophysiological indexes of prediction

There are fewer reports of ERP indexes of the actual prediction as it takes shape before it is confirmed or disconfirmed. However, regarding the variables affecting certainty, Dufour et al. (2013) found more negativity for word onsets of frequent than infrequent words. The effect had a significant, widespread distribution only from 330 ms post word onset. A left-anterior negativity was also visible at 250–330 ms but was not tested statistically in planned comparisons. Due to their hypotheses, the authors primarily investigated frontocentral sites for that time window. The ERPs corresponding to *phonological neighborhood* density, a measure related to lexical competition, have also been assessed (Dufour et al., 2013; Hunter, 2013, 2016; Söderström, Horne, & Roll, 2016). The phonological neighborhood of a word consists of all the words that can be obtained by substituting, adding, or deleting a single phoneme. Sparser neighborhoods, related to lower competition, were observed to increase an ERP negativity between 200 and 300 ms after word onset. Hunter (2013) interpreted the effect as a positive increase for denser neighborhoods. However, in averaged ERPs, it is impossible to distinguish between a positive increase for one condition and a negative expansion for another. Therefore, all things being equal, the effect could be interpreted as a negativity for sparser neighborhoods and thus potentially a reflection of increased certainty about the word ending at word onset.

The contingent negative variation (CNV) is sensitive to anticipation of a future stimulus. It is elicited by the anticipatory association of a sensory stimulus with a subsequent one (Walter et al., 1964). The CNV is thought to reflect expectancy of the second stimulus (S2) upon hearing the first “warning” stimulus (S1). Source localization and fMRI studies have mostly found the supplementary motor area (SMA) in the medial part of the superior frontal gyrus and adjacent cingulate cortex to be the most likely neural sources of the CNV (Gómez et al., 2003; Nagai et al., 2004). Since the SMA is often involved in motor planning, it might be thought that the CNV reflects preparation for a motor response to the second stimulus. However, the CNV has been obtained even without any task, still with the SMA as the most likely source (Mento et al., 2013).

4.1. The pre-activation negativity

The pre-activation negativity (PrAN) (Figure 1) is an electrically negative ERP effect occurring mainly at left-frontal sites of the head. Time-wise, the PrAN overlaps with the P2 and later components, usually starting at 136 ms from word or F0 onset and lasting at least until 280 ms. Two phases of PrAN can be distinguished based on global field power (GFP) analyses of the ERP signal (Lehmann & Skrandies, 1980) and topographical distribution. The early phase (136–200 ms) has a left posterior distribution. The late phase (200 ms onwards) is frontal with a less pronounced left-lateralization. The PrAN increases when native listeners hear word beginnings with highly predictable continuations (Roll et al., 2015; Söderström, Horne, Frid, et al., 2016). The two language-internal factors lexical competition and word frequency of lexical competitors have both been shown to influence PrAN amplitude in the way that would be expected for PrAN to index pre-activation. Specifically, whereas a *decreased* number of lexical competitors of a word-initial diphone enhances PrAN amplitude, *increased* word frequency of the competitors leads to higher amplitude. This relation has been summarized in the linear model in equation (1), adapted from Roll et al. (2017). The constants k and m represent weights of the different terms. The effect has been registered using different behavioral tasks, including acceptability judgments (Roll & Horne, 2011; Roll et al., 2009,

2010, 2011), judging whether a word is in singular/plural or present/past tense form (Hed et al., 2019; Hjortdal et al., 2022; Novén, 2021; Roll, 2015; Roll et al., 2013; Roll et al., 2015; Söderström, Horne, Mannfolk, et al., 2017; Söderström, Horne, & Roll, 2017), pressing a button as soon as a word ends (Gosselke Berthelsen et al., 2018; Roll et al., 2013; Roll et al., 2015), or making a word order judgment (Söderström et al., 2018).

$$(1) PrAN = k(\text{frequency of lexical competitors}) - m(\text{number of lexical competitors})$$

In other words, PrAN has the characteristics that would be expected for an ERP component indexing pre-activation of linguistic material.

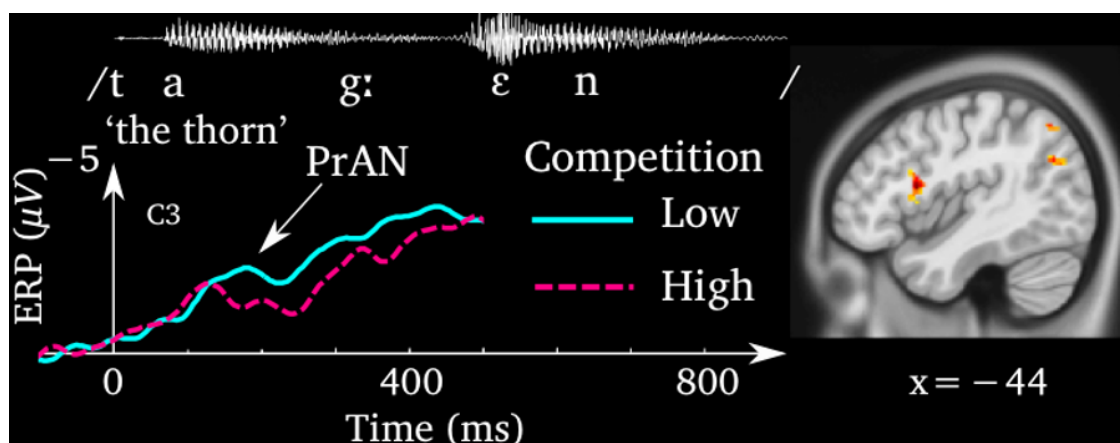


Figure 1. Pre-activation negativity (PrAN) for segmental phonemes (left) at a left-central electrode (C3) and correlated BOLD effect in posterior Broca's area and the left angular gyrus (right). One of the stimulus words (*taggen* 'the thorn') is shown for latency comparison.²

As mentioned above, we cannot observe from an ERP difference whether the effect is an electrical positivity for one condition or a negativity for the other. Thus, how can we know that PrAN is a negative effect for higher certainty about the immediate future and not a positive effect for increased uncertainty about what is coming up? There are two main arguments for PrAN being a negative component. First, GFP analyses of the PrAN signal have shown increased peaks of activity at 136 ms, 200 ms, and 280 ms for word-initial phonemes that yield higher certainty about word endings (Roll et al., 2017; Roll et al., 2015). The peaks show maxima in the electric field strength, argued to indicate the onset of states of brain activity (Khanna et al., 2015). Therefore, the GFP peaks at the beginning of differences between ERPs can be seen as an indication that the neural effect is more likely to happen in the condition where the peak is. In support of this interpretation, in three studies, the GFP difference between high and low prediction conditions at the peak of the GFP for the high prediction condition has been seen to correlate with a significantly increased blood-oxygen-level-dependent (BOLD) signal for the same contrast (Roll et al., 2017; Roll et al., 2015; Söderström, Horne, Mannfolk, et al., 2017; Söderström et al., 2018). The BOLD signal grows when a brain area is put to greater use, indicating that the enhanced negativity in PrAN reflects intensified neuronal activity. PrAN should therefore be considered a neuroelectrically negative effect for lower competition rather than a positive effect for higher competition. This also gives reason for the reinterpretation of Hunter's effect as a negativity for sparser phonological neighborhoods, which is a measure related to decreased lexical competition as described above. Hunter (2016) found that the effects of neighborhood density disappeared when phonotactic probability and cohort size were held constant, indicating that these

² Data from Roll et al. (2017).

measures could be driving the effect. Cohort size is different from the neighborhood definition in that it reflects the number of competitors sharing the first speech sounds. This measure more closely resembles the competition measures used in Roll et al. (2017) and Söderström, Horne, Frid, et al. (2016) which led to negative-going deflections. Taken together, the results indicate that onset effects are likely drivers of PrAN amplitude. This is not surprising since information about rhyme competition is typically not yet available during the first few hundred milliseconds after word onset. Accordingly, in an eye-tracking study, Magnuson et al. (2007) found that effects of onset density emerged before those of neighborhood density. Similarly, phonological neighborhood density effects were only observed in the later PrAN time window, at 208–280 ms, in a study on Swedish (Söderström, Horne, & Roll, 2016). Future studies should control for the effects of phonotactic probability by including variation in both phonotactic probability and lexical competition in the same model.

5. PrAN in response to suprasegmental cues

We propose that PrAN indexes the predictive strength not only of segmental phonemes but of phonemes in general. For the present purposes, we will also include tones with lexical or grammatical associations in the phoneme category. Their phonemic status will be further discussed in section 5.1. In this vein, tonal cues have also been seen to influence PrAN in a similar fashion. Indeed, close scrutiny of the ERP effects of Swedish word accents led to the initial observations of PrAN (Roll et al., 2015; Söderström, Horne, Frid, et al., 2016). PrAN has since been detected in response to the Danish creaky voice feature *stød* (Hjortdal et al., 2022), left-edge boundary tones (Söderström et al., 2018), and, as will be argued below, can also be seen in previous results for right-edge boundary tones (Roll & Horne, 2011).

5.1. Word-level tones

Swedish *word accents* are tonal patterns that are intrinsically tied to the morphological composition of words. In the grammar of Central Swedish speakers, the key features of word accents are the following phonological elements: a low (L*, accent 1) or high (H*, accent 2) tone associated with the stressed syllable of words (Bruce, 1977), which is usually found in the word stem (Figure 2). The word accents are phonologically distinctive, as in the minimal pair ¹*anden* ‘the duck’ and ²*anden* ‘the spirit’, but they have a relatively low functional load in the traditional sense (Elert, 1964). There are only about 350 minimal word accent pairs in Swedish (Elert, 1972), and these differ in terms of word class or morphology (Riad, 2014). Instead, word accents find a more substantial role in their predictive function (Roll, 2022). In addition to the PrAN, the predictive function of word accents is evidenced by increased response times and P600 effects for suffixes that have been invalidly cued by the wrong word accent (Gosselke Berthelsen et al., 2018; Novén, 2021; Roll, 2015; Roll et al., 2010; Roll et al., 2013; Roll et al., 2015; Söderström et al., 2012). Further evidence for their predictive function is the facilitatory effect word accents have in speech processing. Specifically, individuals who give more weight to these tones while listening also process words faster (Roll, 2022).³ This is, to a large extent, explained by the close connection between word accents and morphology outlined below.

³ Roll (2022) tested word-processing speed as a predictor of retardation in responding to suffixes invalidly cued by the wrong word accent. Swapping the variables in a linear regression model in line with the statement above, retardation is a significant predictor of word-processing speed for words with valid word accents, $F(1, 76) = 11.52, p = 0.001$, adjusted $R^2 = 0.120$, but does not predict the response times of words with invalid accents, $F(1, 76) = 1.18, p = 0.281$, adjusted $R^2 = 0.002$. Outliers of $\pm 3 SD$ from the sample mean of the dependent variable were removed.

The most decisive factor for a word’s accent assignment is the suffix: words ending with the singular definite suffix *-en* have accent 1, as in ¹*lek-en* ‘the game,’ whereas words ending with the indefinite plural *-ar* have accent 2, as seen in ²*lek-ar* ‘the games’ (Bruce, 1977; Riad, 2014; Rischel, 1963). Note that although the tone on the stem differs between the two words, the stress is on the first syllable in both cases. Due to the close stem tone-suffix correlation, word accents are excellent predictors of how a word will end. In addition, accent 1 is a much stronger predictor than accent 2. The reason is that accent 1 predicts fewer word continuations than accent 2 (Söderström, Horne, Frid, et al., 2016). Due to a postlexical rule, accent 2 is also assigned to words with secondary stress, regardless of their suffix. Since compound words have secondary stress, all compounds consequently have accent 2. Therefore, when hearing a word beginning with accent 1, listeners can predict a termination in some suffix with relatively high certainty, but when hearing accent 2, a much larger set of lexical competitors opens up. In fact, the lexical competition of word beginnings with accent 2 is 10.5 times larger than that of word beginnings with accent 1, as calculated using the PAROLE corpus (Söderström, Horne, Frid, et al., 2016). This does not mean that accent 2 is more frequent; the two word accents have a similar frequency of occurrence in Swedish.

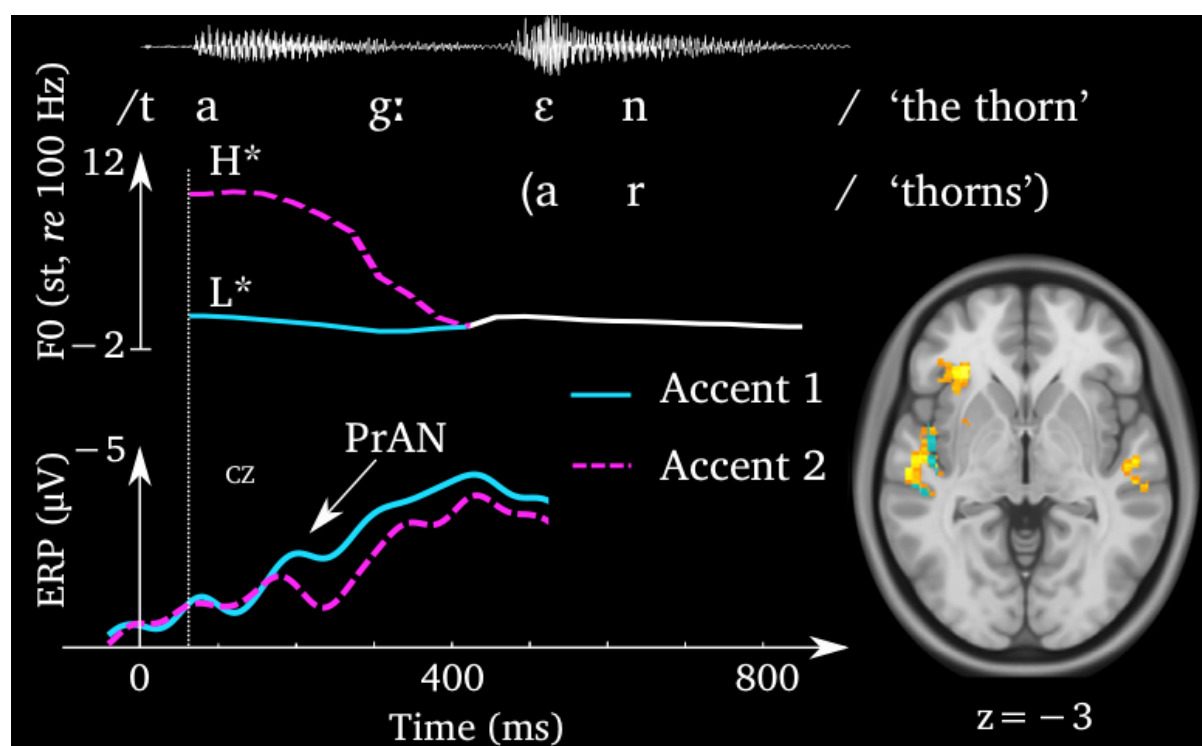


Figure 1. Swedish word accents (top left) and word accent PrAN (bottom left). Correlated BOLD activity in left temporal cortex (right), involving primary and secondary auditory cortices, as well as predominantly anterior Broca’s area.

Accent 1 has been shown in several studies to have a larger negativity than accent 2 during the first 136–300 ms (Roll et al., 2010; Roll et al., 2013; Roll et al., 2015; Söderström, Horne, Mannfolk, et al., 2017; Söderström, Horne, & Roll, 2017). This has been interpreted as a PrAN effect due to the lower number of lexical competitors (Söderström, Horne, Frid, et al., 2016) and the consequent possibilities for increased lateral inhibition of irrelevant word forms (Roll et al., 2017). Notice that the accent 1 PrAN is not a purely acoustic effect but is rather phonologically driven. Thus, for test stimuli consisting of the pitch contour alone (hummed speech created using Praat (Boersma & Weenink, 2001), there was no negativity for accent 1, but instead an N1 effect for accent 2, due to the acoustic salience of the H* tone peak. The H* was followed by a fall of, on average, 7.4 semitones, compared to the L* tone, with, on average, a 1.0 semitone fall (Roll et al., 2013). Likewise, in South Swedish, where word

accents are practically the tonal mirror image of Central Swedish accents but are nevertheless functionally similar, accent 1 still produced a PrAN (Roll, 2015). As with the segmental PrAN, GFP effects and correlated BOLD increase have indicated that the negativity for accent 1 is associated with augmented neural activity (Roll et al., 2015). Lastly, learners who had not yet acquired the predictive function of word accents did not show an accent 1 PrAN (Gosselke Berthelsen et al., 2018). However, a generally increased PrAN and a PrAN differentiation between word accents developed after intense phonological training (Hed et al., 2019).

5.2. Clause-level tones

Not only has PrAN been observed at the word level, but also at the syntactic level, where tonal cues to syntactic structure produce increased negativity. Thus, the Central Swedish “left-edge boundary tone” (Roll, 2006; Roll et al., 2009) or “initiality accent” (Myrberg, 2010) is a high tone in the last syllable of the first prosodic word of main clauses. It does not, however, occur in subordinate clauses (Roll, 2006). Therefore, the presence or absence of a left-edge boundary tone is a good predictor of the syntactic structure of a clause. Clauses beginning with the subordinate conjunction *att* ‘that’ can have either subordinate or (embedded) main clause structure. The structure is disambiguated in the presence of sentence adverbs like the negator *inte* ‘not,’ which follow the inflected verb in main clauses (...*att Gunnar kommer inte* ‘that Gunnar comes not’) but precedes the verb in subordinate clauses (...*att Gunnar inte kommer* ‘that Gunnar not comes’) (Holmberg & Platzack, 1995). In short, with main-clause structure comes a high tone on the last syllable of the first prosodic word: *att Gunnar^H kommer inte* ‘that Gunnar^H comes not.’ Listeners use the presence/absence of a left-edge boundary tone to predict the clause structure, as shown by structural reanalysis (updating) effects (P600) in case of tone-word order mismatch (Roll & Horne, 2011; Roll et al., 2009, 2011; Söderström et al., 2018). Since main clauses involve a larger set of structural options (different types of topicalization and force) unavailable to subordinate clauses, the *absence* of a tone is the best structural predictor in *att* ‘that’ clauses. Accordingly, the absence of a left-edge boundary tone at the beginning of these clauses has been observed to produce increased negativity (Roll et al., 2009, 2011), which has been interpreted as a PrAN (Söderström et al., 2018). This negativity has also been found to correspond to larger GFP and BOLD effects, indicating a relation to increased neuronal activity (Söderström et al., 2018).

The presence of *right*-edge boundary tones, marking the end of intonation phrases, is also a good predictor of syntactic structure. In Swedish, right-edge boundary tones are usually low (L%) (Bruce, 1977). In this vein, Roll and Horne (2011) used sentences with or without (Ø) right-edge (L%) and left-edge (H) boundary tones like *Sheriffen bakband boven^{L%/Ø} och botanikern^{H/Ø} strök/stramt...* ‘The sheriff tied the villain^{L%/Ø} and the botanist^{H/Ø} prowled/tightly...’ to investigate the effects of the interaction of boundary markers during online listening (Figure 3). At the noun phrase *botanikern* ‘the botanist,’ the sentence is structurally ambiguous. *Botanikern* ‘the botanist’ might belong to a continuation of the first clause, so that the string *boven och botanikern* ‘the villain and the botanist’ forms a coordinated object noun phrase. In that case, an adverb like *stramt* ‘tightly’ might follow. However, *botanikern* ‘the botanist’ could also begin a new, coordinated main clause. In that case, it could only be followed by a verb like *strök* ‘prowled.’ Clause continuation/noun phrase coordination is compatible with the absence of right- and left-edge boundary tones. A new main clause, on the other hand, requires the presence of both boundary tones.

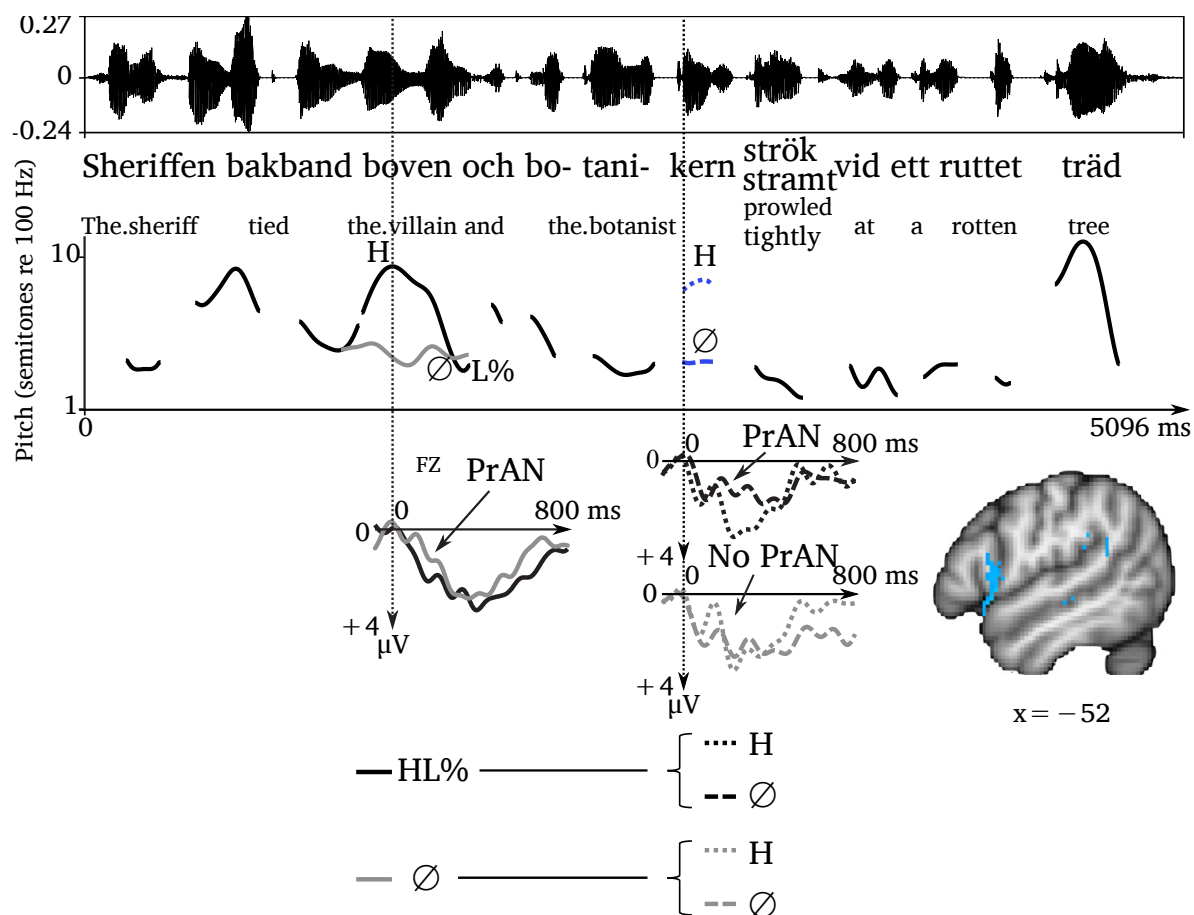


Figure 2. Two different sentence structures associated with the presence or absence (\emptyset) of right-edge (L%) or left-edge (H) boundary tones. There is a pre-activation negativity (PrAN) for the absence of a right-edge boundary tone (\emptyset) as compared to its presence (L%) on *boven* ‘the villain,’ since its absence cues sentence continuation (more restrictive). A PrAN is further seen for the absence of a left-edge boundary tone on *botanikern* ‘the botanist,’ also cueing sentence continuation, but only after the presence of a preceding right-edge boundary tone, where the structural possibilities are still more open. The PrAN for the absence of a left-edge boundary tone, cueing clause continuation, has been related to activation in Broca’s area, as shown on the brain to the right.⁴

As in the case of the left-edge boundary tone, the *absence* of a right-edge boundary tone is a better structural predictor than its presence. Thus, before producing a new intonation phrase containing a new main clause, the previous clause and intonation phrase need to be closed. Therefore, if hearing *Sheriffen bakband boven^Ø och...* ‘The sheriff tied the villain^Ø and...’ *without* a right-edge boundary tone, the listener can be sure that the ongoing clause/intonation phrase will continue. In this context, the following noun phrase *botanikern* ‘the botanist’ will be predicted to be part of a coordinated object noun phrase *boven och botanikern* ‘the villain and the botanist,’ and not a constituent of a new clause. However, if *Sheriffen bakband boven^{L%} och...* ‘The sheriff tied the villain and...’ is produced with a right-edge boundary tone on *boven* ‘the villain,’ a new clause is predicted to start, which means a greater degree of structural uncertainty. Thus, *botanikern* ‘the botanist’ might be a subject or a topicalized object or form part of some larger constituent. Similar to what happens with word-level PrAN, the more predictive condition, the absence of a right-edge boundary tone,

⁴ Data from Roll & Horne (2011) and Söderström et al. (2018).

produced a PrAN-like negativity between 100-250 ms.⁵

The absence of a right-edge boundary tone is such a strong predictor of sentence continuation that a following missing boundary tone on *botanikern* ‘the botanist’ in *Sheriffen bakband boven⁰ och botanikern⁰* becomes less informative. Therefore, if there was no right-edge boundary tone in the preceding noun phrase, the absence of left-edge boundary produced no PrAN. If anything, there was a rapid negativity for the *presence* of a left-edge boundary tone in *Sheriffen bakband boven⁰ och botanikern^H...* ‘The sheriff tied the villain⁰ and the botanist^H...’ However, the latency of this negativity is too early (50-150 ms) for it to be interpreted as a syntactic PrAN. It might rather be, along with the interpretation of the authors, an N1 effect showing increased prediction error due to the occurrence of a highly improbable tone. Following a right-edge boundary tone, on the other hand, as in *Sheriffen bakband boven^{L%} och...* ‘The sheriff tied the villain^{L%} and...’ the clause is expected to end at *boven^{L%}* ‘the villain^{L%},’ and a new clause is expected to start at *botanikern* ‘the botanist.’ In this context, the absence of a left-edge boundary tone in *botanikern⁰* ‘the botanist’ changes the expectation, indicating sentence continuation with a coordinated noun phrase (*boven och botanikern* ‘the villain and the botanist’). The absence of a tone is thus informative and increases expectation for the option with fewer possible continuations, that is, sentence continuation rather than a new clause. It thus leads to increased structural certainty and, hence, increased PrAN compared to the presence of a tone, as usual. To sum up, tonal environments leading to increased certainty of the continuation within or between words increase PrAN in the same way that segments do.

5.3. Phonological and phonetic information

As noted above, the PrAN seems to reflect phonological function, rather than phonetic or acoustic processing, as evidenced by PrAN effects for both Central (Roll et al., 2010; Roll et al., 2013; Roll et al., 2015) and South Swedish accent 1 (Roll, 2015), despite the pitch realizations being, to some extent, each other’s mirror images. Recently, the effects of phonetic and phonological cues were dissociated in a study with the Danish creaky voice feature ‘stød’ and its modal voice counterpart ‘non-stød’ (Hjortdal et al., 2022). Stød is genetically related to Swedish accent 1 but is phonetically very distinct. Stød is often described as having two phases (Fischer-Jørgensen, 1989). Phase 1 shows differences in pitch that covary with phase 2 (Peña, 2022). Phase 2 consists mainly of a creaky voice realization, which has been considered the phonological locus of stød (Basbøll, 2014; Fischer-Jørgensen, 1989). Like Swedish word accents, stød and non-stød covary with different suffixes and can be used as suffix predictors in speech perception. Also similar to accent 1, stød is around four times as predictively useful as non-stød. Hjortdal et al. (2022) spliced stimuli so that stød phase 1, stød phase 2, and suffixes were crossed. The validity of both phases influenced the response times, although phase 2 more so than phase 1. Stød phase 2, as compared to non-stød, resulted in an anterior negativity between 280 and 430 ms, interpreted as a late PrAN. The phonetic cues to stød during phase 1 did not result in increased PrAN amplitude. The stronger predictive value of the phonological phase 2 was also seen in the fact that suffixes mismatching preceding stød or non-stød phase 2 cues yielded N400 and P600 effects. This was not the case for suffixes mismatching phase 1. Bornkessel-Schlesewsky & Schlesewsky (2019) have proposed that updating of the internal generative model, as reflected in the N400 amplitude, is modulated by the availability and reliability of linguistic cues. Phonological

⁵ This is a re-interpretation of the effect, since at the time, PrAN had not been discovered, and therefore, the effect was interpreted as a positivity for the presence of a right-edge boundary tone, without measuring GFP. Interestingly, however, in favor of the PrAN hypothesis, a negative deflection similar to those found in other PrAN studies can be observed in the ERPs.

cues might be stronger predictors than phonetic covariation since they are more invariant and thus more reliable.

6. PrAN and other ERP components

Although MMN, N1, ELAN, and the early PMN all bear superficial similarities to the first phase of PrAN (136-200 ms), there is a major difference: whereas PrAN indexes the prediction (feed-forward) process, the other components reflect some aspect of prediction error (feedback process). The distinction is transparent in the paradigms eliciting the different effects. ELAN, PMN, and MMN paradigms create expectations of various kinds: for a grammatical morpheme through a phrase structure context (ELAN), for a specific word form by different means (PMN), or for a word form or morpheme by repeated presentation of one stimulus (MMN). The expectations are then mismatched by some stimulus, producing the neural result. N1 effects are usually also elicited by broken expectations and, importantly, increased attention allocation, as described above. The PrAN paradigms, on the other hand, present phonological cues in neutral contexts, where expectations are constant between conditions at cue onset. The variation in the signal is obtained by the predictive potential of the cue itself, not by varying its context.

In terms of timing, PrAN occurs after N1, and its first phase rather overlaps with the P2 component, leading early studies to report PrAN as a P2 modulation. However, although PrAN temporally coincides with the positive P2 component, it corresponds to increased *negativity* for predictively useful phonological cues, as evidenced by the GFP and BOLD contrast correlation reported above. Roll et al. (2013) dissociated PrAN from the N1 both temporally and functionally. When participants listened to isolated words, accent 1 produced a PrAN compared to accent 2. The negativity overlapped with the P2 component in early (150–200 ms) and later (200–300 ms) time windows (Roll et al., 2013). The effect was visible in the upstroke of an unequivocal P2 component. No difference between word accents was detected during the likewise prominent N1 component. Conversely, when the speech melody was presented in delexicalized stimuli, containing only the F0 contour, the N1 (100–150 ms) increased for the acoustically more salient high accent-2 tone compared to the low accent-1 tone. This difference was observed to extend over a noticeable N1-component downstroke, with no effect in the P2 time range. The results are in line with an interpretation of the N1 as showing increased prediction error for sounds that are unexpected due to the context or their auditory salience (Astheimer & Sanders, 2011; Hillyard et al., 1973; Nobre & van Ede, 2018; Roll & Horne, 2011; Roll et al., 2013). The gain of the N1 effect can be modulated in relation to the relevance of a temporal position, making it increase, for example, for word onsets (Astheimer & Sanders, 2009; Sanders & Neville, 2003). With the evidence at hand, it is difficult to say how much attention affects the PrAN. As mentioned above, the effect has been obtained using different tasks. However, to date, no study has investigated whether a PrAN is obtained in the absence of a task.

Regarding its latency, PrAN is similar to the MMN. However, as argued above, while the MMN is sensitive to the physical characteristics of stimuli or how unexpected they are in a certain context, PrAN reflects the stimuli's predictive potential. For example, as discussed above, PrAN is greater for accent 1 than for accent 2 due to it being a stronger predictor for word endings irrespective of whether accent 1 is realized as a low (Roll et al., 2015; Söderström, Horne, & Roll, 2017) or a high tone (Roll, 2015), and even if both word accents can be equally expected based on their frequency. PrAN also disappears for tonal contrasts in the absence of segmental content (Roll et al., 2013). The most commonly reported PMN time window overlaps with part of the second phase of PrAN. Like MMN, the PMN is also related to prediction error—increasing for unexpected sounds—rather than prediction, which would mean an increase for higher certainty. The PMN is further found for unexpected speech sounds regardless of lexicality. Segmental PrAN, on the other hand, is difficult to define for non-

existing word beginnings since it has, so far, been measured in terms of lexical competition and frequency of competitors, measures that are inherently absent for pseudowords. A PrAN has been shown for word accents in pseudowords (Söderström, Horne, Mannfolk, et al., 2017; Söderström, Horne, & Roll, 2017). However, the pseudowords included real suffixes that the word accents were associated with. The association between word accents and suffixes was so strong that participants could even recover the meaning of ~80% of suffixes masked by coughs using only the word accent information (Söderström, Horne, & Roll, 2017).

Functionally, PrAN shares many characteristics with the CNV. The experiment design with one stimulus cueing another is relatively similar to the conditions under which PrAN is found. In this case, the phonemes at the beginning of a word cue different possible word endings. Further, the CNV for a stimulus 1 (S1) has greater amplitude for a highly probable stimulus 2 (S2) than for an S2 with a lower probability. This is similar to what has been observed for PrAN, where the amplitude increases the more constrained the possibilities are for word endings. Although there are similarities, there are three main differences between CNV and PrAN, however. First, whereas PrAN is observed as early as 136 ms following stimulus onset, CNV is typically calculated from 280 ms from S1 onwards. Second, whereas CNV is typically rather evenly distributed over central electrodes, PrAN has had a clear left-lateralized distribution, more typical of language processing (Shtyrov et al., 2005). Third, PrAN is seen in response to language, a form of “overlearned” sound-sensory-motor associations. The late timing and frontal sources of CNV are similar to the later phase of PrAN. Considering the general inhibitory and disinhibitory function of frontal lobe structures (Rocchetta & Milner, 1993; Sumner et al., 2007), it could be that both index predictive processing by inhibition of irrelevant alternatives and disinhibition of relevant alternatives: words or syntactic structure in the case of PrAN, and most often spatial locations when it comes to CNV. A recently found CNV-like anterior negativity developing over semantically increasingly constraining sentences supports this hypothesis (Grisoni et al., 2017; Grisoni et al., 2021; León-Cabrera et al., 2019; León-Cabrera et al., 2017). Thus, similar to the late PrAN, Grisoni et al. (2021) found probable sources for their negativity in more predictive contexts in the inferior frontal gyrus.

7. Possible brain sources

Recent neurolinguistic models assume two different streams of language processing in the brain, the *dorsal* and *ventral* streams (Hickok & Poeppel, 2004; Saur et al., 2008). Both streams start in the primary auditory cortex in Heschl’s gyrus, situated at the hidden surface inside the superior temporal gyrus. Both also pass through what can be described as the secondary auditory cortex, the planum temporale, which is found lateral and posterior to Heschl’s gyrus (DeWitt & Rauschecker, 2013). The dorsal stream then connects to the frontal lobe through the parietal lobe and superiorly located pathways, whereas the ventral stream goes anteriorly through the superior and middle temporal lobe and connects to the frontal cortex through inferior pathways. Whereas the ventral stream is involved in automatically connecting word forms to meaning, the dorsal stream is more involved in auditory-motor mapping and syntactic processing (Friederici, 2017). The sound-articulation connection in the dorsal stream is relevant for language learning, where repetition is important (Hickok & Poeppel, 2004) as well as phonetic and phonological memory (Kellmeyer et al., 2013; Novén, Olsson, et al., 2021; Saur et al., 2008). This connection is also probably what makes the dorsal stream more active during effortful listening under noisy conditions (Garrod et al., 2014). The dorsal stream is also more involved in syntactic (Skeide et al., 2016) and decompositional morphological processing (Schremm et al., 2018).

Prediction is thought to mediate processing in both streams. However, it can be assumed to be of varying nature depending on the stream. Thus, the ventral stream would involve automatic pre-activation from a higher area to a lower (Hickok, 2012). For example, expected

word forms might pre-activate the upcoming phonemes they contain. In the other direction, the phonemes actually encountered would trigger a prediction error from the lower phoneme-processing area to the higher word-processing area. This kind of prediction is difficult to distinguish psycholinguistically from a bottom-up model. Prediction in the dorsal pathway is easier to grasp since it involves auditory-motor connections. Hence, it is what we can feel we are doing during effortful listening, trying to articulate what we think we hear (Garrod et al., 2014). This can occur to different degrees. Lower degrees of pre-activation involving the dorsal stream are probably prevalent. The degree of involvement is likely to increase with listening effort until reaching half-conscious articulation at the extreme end.

Speech processing before 200 ms is thought to involve “bottom-up” processing in the ventral stream (Skeide & Friederici, 2016), through automatic hierarchical prediction in the present framework. The timing, spatial distribution, and possible sources coincide well with the characteristics of early PrAN (136–200 ms). This effect has had a left posterior distribution and has correlated with BOLD effects in Heschl’s gyrus (primary auditory cortex), the superior temporal gyrus (secondary auditory cortex), and the left inferior frontal gyrus (IFG), pars orbitalis (the anterior portion of Broca’s area, Brodmann area (BA) 47). These areas form part of the ventral processing stream (Friederici et al., 2017; Hickok & Poeppel, 2004). BOLD correlates for PrAN in the primary and secondary auditory cortex have only been found using the word accent PrAN contrast (Roll et al., 2015; Söderström, Horne, Mannfolk, et al., 2017). Even if no BOLD correlations have been reported in this early time frame for segmental PrAN, it also shows a similar left-central-to-posterior negativity that is clearly predictively loaded since it correlates with reduced lexical competition and increased frequency of the competitors (Roll et al., 2017). In other words, early PrAN would be thought to show automatic predictive processing in the ventral stream. Adding to this interpretation, Schremm et al. (2018) discovered that cortical thickness of the planum temporale, comprising secondary auditory cortex and forming part of both the ventral and the dorsal stream (DeWitt & Rauschecker, 2013), correlated with response times in judging whether real words were in singular or plural form. Specifically, a thicker planum temporale correlated with a relative increase in response times for incorrect word accent-suffix combinations and generally faster judgments of words with correct word accent-suffix combinations. In short, a thicker planum temporale was related to greater use of word accents as suffix predictors in real words. The authors interpreted the results as an association between thicker cortex and more robust full-form representations of real words and their associated suffixes in the ventral stream. Novén, Schremm, et al. (2021) presented further evidence for this interpretation in a similar correlation between cortical thickness and response times to word accents in a more anterior part of the ventral stream in the temporal lobe.

From around 200 ms after word onset, the processing is thought to involve the dorsal stream and top-down processes to a greater degree (Skeide & Friederici, 2016). This is in line with the time frame, topography, and BOLD correlates of late PrAN. Thus at 200 ms, segmentally induced PrAN correlated with BOLD effects in IFG, pars opercularis (posterior portion of Broca’s area, BA44), and the inferior parietal lobe (angular gyrus, BA39). At 256 ms (Roll et al., 2015) and 320 ms (Söderström, Horne, Mannfolk, et al., 2017), PrAN for predictive word accents was seen to correlate with BOLD activity in the IFG, pars orbitalis (BA47) and opercularis (BA44), respectively. At 220 ms, left-edge boundary-elicited PrAN correlated with BOLD in IFG, pars opercularis (BA44) (Söderström et al., 2018). The activity of mainly posterior Broca’s area and the inferior parietal lobe supports the involvement of the dorsal stream in late PrAN. When interpreting these results, it should be kept in mind that the posterior portion of Broca’s area is known not only as part of the dorsal stream but also as the main locus of syntactic processing (Friederici et al., 2017). The syntactic function of posterior Broca’s area might be strongly associated with its phonological function. Thus, phonologically, this area could be involved in suppressing word forms outside the set of lexical competitors. In a similar way, syntactically, it may be engaged in inhibiting irrelevant

clause structures. The frontmost area of the dorsal stream, IFG, pars opercularis (BA44), indeed seems to be important for using word accents as predictive cues in pseudowords, which by definition do not have any full form representations (Söderström, Horne, Mannfolk, et al., 2017). Some of the areas where BOLD has been found to correlate with late PrAN have also previously been associated with spatial orientation of acoustic attention, which also involves inhibiting and disinhibiting receptive fields in the auditory cortex (Alho et al., 2015). Interestingly, as mentioned above, sources of the late PMN, with a time frame overlapping that of late PrAN, have similarly been found in left anterior cortical regions (Connolly et al., 2001; Kujala et al., 2004).

Finally, the activation in primary auditory cortex for the more predictively useful accent 1 as compared to the acoustically more salient high-toned accent 2 might be interpreted as lending support to the predictive coding framework: activity in primary auditory cortex corresponds to predictivity rather than to sound salience (Gagnepain et al., 2012). Thus, it shows signs of indexing pre-activation rather than bottom-up processing. The fact that the primary auditory cortex activity is found only for the word accent PrAN, but not for the segmental or syntactic PrANs, which show more focus on later left frontal activation, can be explained by the massive reduction of lexical competition that accent 1 entails. As noted above, since accent 1, on average, cues 10.5 times fewer word endings than accent 2 (Söderström, Horne, Frid, et al., 2016), the listener's certainty becomes radically higher when hearing accent 1. The strong association of accent 1 with a limited number of word endings might favor low-level pre-activation of the phonological form of those endings in the primary auditory cortex. Segmental cues to full word forms do not, on average, differ as drastically as word accents in their predictive power and, therefore, might not produce as strong differences in terms of low-level pre-activation. Tonal cues to syntactic structure would not be expected to pre-activate any specific word forms at all, but rather, abstract syntactic patterns. A similar tendency might be discerned for the PMN. As mentioned above, studies finding an early, posterior PMN have involved very specific phonological expectations with specific mismatches. In the D'Arcy et al. (2000) experiment, the expectation for a word form corresponding to a specific color or shape was strong. The early-PMN-producing mismatching terms were also of the same reduced category. Hence, their onsets could potentially lead to a likewise streamlined reweighting of the acoustic prediction. In van den Brink et al. (2001)'s experiment, the expectation can also be assumed to have been extremely specified due to the strong cloze probability.

To sum up, the PrAN component would appear to be associated with two time windows with slightly differing sources. At the first stage, ~136-200 ms, it reflects increased pre-activation in the primary auditory cortex and surrounding areas through automatic hierarchical prediction in the ventral stream. At a later stage, after 200 ms, PrAN is associated with the inhibition of irrelevant candidates, be it word forms, morphemes, or syntactic structures, engaging Broca's area. For boundary tones, this process involves the posterior part of Broca's area (IFG, pars opercularis, BA44) and the connected dorsal stream, activating expected upcoming syntactic structures. The syntactic structures are abstract and do not involve activation of specific phonemes or even word forms, and thus do not engage the primary auditory cortex. Word accents, particularly accent 1, can give rise to much more specific pre-activations of word endings and their auditory representations in the primary auditory cortex. Word accents in frequent nouns are thought to be stored both together with their full word forms along the ventral stream and with their associated suffixes, involving combinatorial mechanisms in the dorsal stream. Word accents can therefore trigger pre-activation in both the ventral and dorsal streams. The ventral stream, however, seems to be most strongly engaged in frequent noun processing, triggering activity in the anterior part of Broca's area (BA45) and the rostrally located pars orbitalis of the inferior frontal gyrus (BA47) in addition to temporal lobe regions, areas which are connected with the ventral stream, and only more weakly with the posterior part of Broca's area (BA44) (Roll et al., 2015). The dorsal

stream, and particularly BA44, is, however, important for processing pseudowords with real suffixes, which do not have any full-form representations in the ventral stream (Schremm et al., 2018). Segments can also be thought to involve prediction in both streams. However, so far, evidence has only been found for segmentally based PrAN sources in areas along the dorsal stream (BA44 and the angular gyrus of the parietal lobe). The lack of visible activation in the primary auditory cortex for segmental PrAN might be due to the lexical competition being too extensive for most word-initial diphones, giving rise to too unspecific and short-lived pre-activations for BOLD increase to be significantly measured.

8. Prediction and error—shared neurobiological substrates?

A final question is how pre-activation can be represented at the neural level. The signal measured using ERPs is likely to show prediction error predominantly. Strictly speaking, it might even be that prediction error and belief updating are the *only* processes that can be measured using ERPs (Schröger et al., 2015). As mentioned above, N1, ELAN, MMN, and PMN can be thought to show prediction error and, perhaps, belief updating (Friston et al., 2021). The question is which neural substrates can be associated with components argued to index predictive content, like CNV or PrAN. A neural process that is likely to be involved in hierarchical prediction during speech processing is the regulation of the inhibition and disinhibition of auditory neurons by interneurons (Almeida, 2021; Chen et al., 2015; Garrett et al., 2020). As mentioned above, auditory cortices are probably involved in generating early PrAN. Thus, when a predictively useful phonological cue arrives at the perceptual system, the hypotheses entertained by the auditory cortex become reweighted. This involves disinhibiting the neurons of the receptive field used for processing features involved in the new dominant hypotheses while increasing the inhibition of other neurons. In terms of predictive coding, this process involves prediction error and belief updating. However, above all, it indexes a form of pre-activation, since it reflects a reweighting of the predictive hypotheses. A similar reweighting is likely to occur in higher-level areas in frontal cortex, being, however, only indirectly related to the low-level auditory predictions, as previously argued for the N400 (Almeida, 2021). In this sense, it might be that predictive and prediction error-indexing potentials partly share their neural substrates. When hearing an unexpected phoneme, the predictive hypotheses are mismatched and reweighted. Assuming that perception is based on a constant predictive weighting of the inhibition of neurons, even in a neutral context, a predictively informative phoneme will produce a reweighting of neural inhibitions. In this sense, although their paradigms and targeted brain areas differ, at the neurobiological level, the ELAN, LAN, MMN, and PrAN might reflect similar processes.

9. Conclusions

The ERP component “pre-activation negativity” (PrAN) indexes the predictive strength of phonological cues. In other words, the better predictor a segment or tone is, the higher its PrAN amplitude is. Specifically, all segments or word-level tones give cues about the ending of unfolding words during online speech perception. These phonological cues can have different degrees of predictive power, and thus influence the PrAN differently; the more certainty they produce about the word completion, the more they boost the PrAN. Similarly, at the sentence level, tones cueing fewer possible upcoming structures also increase the certainty about what is to come and therefore produce a greater PrAN. PrAN can be divided into two phases, one early (136–200 ms) and one late (200 ms onwards). The early phase seems to reflect increased activity, especially in the primary auditory cortex and surrounding regions. Increased activity in the auditory cortex for more predictively useful sounds rather than more salient sounds speaks in favor of a predictive coding framework, where word-level predictions increase activity in the neurons that will process the expected auditory

information. The later PrAN phase correlates more with sources in Broca's area in the frontal lobe, especially its posterior part. This probably reflects the disinhibition of relevant information and inhibition of irrelevant information outside the set of lexical competitors or the set of competing syntactic structures.

Funding information

This work was supported by the Swedish Research Council (Grants No. 2018.00632 and 2019.03063), Knut and Alice Wallenberg Foundation (Grant No. 2018.0454), and Marcus and Amalia Wallenberg Foundation (Grant No. 2018.0021).

References

- Alho, K., Salmia, J., Koistinena, S., Salonene, O., & Rinnea, T. (2015). Top-down controlled and bottom-up triggered orienting of auditory attention to pitch activate overlapping brain networks. *Brain Research*, *1626*, 136-145.
<https://doi.org/10.1016/j.brainres.2014.12.050>
- Almeida, V. N. (2021). Neurophysiological basis of the N400 deflection, from mismatch negativity to semantic prediction potentials and late positive components. *International Journal of Psychophysiology*.
<https://doi.org/10.1016/j.ijpsycho.2021.06.001>
- Andersen, G. (2011). *Leksikalsk database for svensk*.
- Astheimer, L. B., & Sanders, L. D. (2009). Listeners modulate temporally selective attention during natural speech processing. *Biological Psychology*, *80*(1), 23-34.
<https://doi.org/10.1016/j.biopsycho.2008.01.015>
- Astheimer, L. B., & Sanders, L. D. (2011). Predictability affects early perceptual processing of word onsets in continuous speech. *Neuropsychologia*, *49*(12), 3512–3516.
<https://doi.org/10.1016/j.neuropsychologia.2011.08.014>
- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445-459. <https://doi.org/10.3758/bf03193014>
- Basbøll, H. (2014). Danish stød as evidence for grammaticalization of suffixal positions in word structure. *Acta Linguistica Hafniensia*, *46*(2), 137-158.
<https://doi.org/10.1080/03740463.2014.989710>
- Boersma, P., & Weenink, D. (2001). PRAAT, a system for doing phonetics by computer. *Glott International*, *5*, 341-345.
- Bornkessel-Schlesewsky, I., & Schlewsky, M. (2019). Toward a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in Psychology*, *10*(298), 1-17. <https://doi.org/10.3389/fpsyg.2019.00298>
- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Gleerups.
- Chen, I.-W., Helmchen, F., & Lütcke, H. (2015). Specific early and late oddball-evoked responses in excitatory and inhibitory neurons of mouse auditory cortex. *Journal of*

- Neuroscience*, 35(36), 12560-12573. <https://doi.org/10.1523/jneurosci.2240-15.2015>
- Cohen, C. (2014). Probabilistic reduction and probabilistic enhancement. *Morphology*, 24(4), 291-323. <https://doi.org/10.1007/s11525-014-9243-y>
- Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, 6(3), 256-266. <https://doi.org/10.1162/jocn.1994.6.3.256>
- Connolly, J. F., Phillips, N. A., Stewart, S. H., & Brake, W. G. (1992). Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences. *Brain and Language*, 43(1), 1-18. [https://doi.org/10.1016/0093-934x\(92\)90018-a](https://doi.org/10.1016/0093-934x(92)90018-a)
- Connolly, J. F., Service, E., D'Arcy, R. C. N., Kujala, A., & Alho, K. (2001). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping. *Neuroreport*, 12(2), 237-243. <https://doi.org/10.1097/00001756-200102120-00012>
- Connolly, J. F., Stewart, S. H., & Phillips, N. A. (1990). The effects of processing requirements on neurophysiological responses to spoken sentences. *Brain and Language*, 39, 302-318. [https://doi.org/10.1016/0093-934x\(90\)90016-a](https://doi.org/10.1016/0093-934x(90)90016-a)
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology-Human Perception and Performance*, 14(1), 113-121. <https://doi.org/10.1037/0096-1523.14.1.113>
- D'Arcy, R. C. N., Connolly, J. F., & Crocker, S. F. (2000). Latency shifts in the N2b component track phonological deviations in spoken words. *Clinical Neurophysiology*, 111(1), 40-44. [https://doi.org/10.1016/S1388-2457\(99\)00210-2](https://doi.org/10.1016/S1388-2457(99)00210-2)
- Davis, H. (1964). Enhancement of evoked cortical potentials in humans related to a task requiring a decision. *Science*, 145(3628), 182-183. <https://doi.org/10.1126/science.145.3628.182>
- DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117-1121. <https://doi.org/10.1038/nn1504>
- DeWitt, I., & Rauschecker, J. P. (2013). Wernicke's area revisited: Parallel streams and word processing. *Brain and Language*, 127(2), 181-191. <https://doi.org/10.1016/j.bandl.2013.09.014>
- Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11(3), 357-374. <https://doi.org/10.1017/S0140525X00058027>
- Dufour, S., Brunellière, A., & Frauenfelder, U. H. (2013). Tracking the time course of word-frequency effects in auditory word recognition with event-related potentials. *Cognitive Science*, 34, 489-507. <https://doi.org/10.1111/cogs.12015>

- Duta, M. D., Styles, S. J., & Plunkett, K. (2012). ERP correlates of unexpected word forms in a picture–word study of infants and adults. *Developmental Cognitive Neuroscience*, 2(2), 223-234. <https://doi.org/10.1016/j.dcn.2012.01.003>
- Elert, C.-C. (1964). *Phonologic studies of quantity in Swedish based on material from Stockholm speakers*. Almqvist & Wiksell.
- Elert, C.-C. (1972). Tonality in Swedish: Rules and a list of minimal pairs. In E. S. Firchow, K. Grimstad, N. Hasselmo, & W. O’Neil (Eds.), *Studies for Einar Hagen*. Mouton.
- Ettinger, A., Linzen, T., & Marantz, A. (2014). The role of morphology in phoneme prediction: evidence from MEG. *Brain and Language*, 129, 14-23. <https://doi.org/10.1016/j.bandl.2013.11.004>
- 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>
- Fischer-Jørgensen, E. (1989). Phonetic analysis of the stød in standard Danish. *Phonetica*, 46(1-3), 1-59. <https://doi.org/10.1159/000261828>
- Fitzroy, A. B., & Sanders, L. D. (2015). Musical meter modulates the allocation of attention across time. *Journal of Cognitive Neuroscience*, 27(12), 2339-2351. <https://doi.org/10.1162/jocn.a.00862>
- Fitzroy, A. B., & Sanders, L. D. (2021). Subjective metric organization directs the allocation of attention across time. *Auditory Perception & Cognition*, 3(4), 212-237. <https://doi.org/10.1080/25742442.2021.1898924>
- Friederici, A. D. (2017). *Language in our brain*. MIT Press.
- Friederici, A. D., Chomsky, N., Berwick, R. C., Moro, A., & Bolhuis, J. J. (2017). Language, mind and brain. *Nature Human Behaviour*, 1(10), 713-722. <https://doi.org/10.1038/s41562-017-0184-4>
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 815-836. <https://doi.org/10.1098/rstb.2005.1622>
- Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293-301. <https://www.ncbi.nlm.nih.gov/pubmed/19559644>
- Friston, K. (2018). Does predictive coding have a future? *Nature Neuroscience*, 21, 1019-1026. <https://doi.org/10.1038/s41593-018-0200-7>
- Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 1211-1221. <https://doi.org/10.1098/rstb.2008.0300>
- Friston, K. J., Sajid, N., Quiroga-Martinez, D. R., Parr, T., Price, C. J., & Holmes, E. (2021). Active listening. *Hearing Research*, 399, 107998. <https://doi.org/10.1016/j.heares.2020.107998>

- Gagnepain, P., Henson, Richard N., & Davis, Matthew H. (2012). Temporal predictive codes for spoken words in auditory cortex. *Current Biology*, 22(7), 615-621.
<https://doi.org/10.1016/j.cub.2012.02.015>
- Garrett, M., Manavi, S., Roll, K., Ollerenshaw, D. R., Groblewski, P. A., Ponvert, N. D., Kiggins, J. T., Casal, L., Mace, K., Williford, A., Leon, A., Jia, X., Ledochowitsch, P., Buice, M. A., Wakeman, W., Mihalas, S., & Olsen, S. R. (2020). Experience shapes activity dynamics and stimulus coding of VIP inhibitory cells. *eLife*, 9, e50340.
<https://doi.org/10.7554/eLife.50340>
- Garrod, S., Gambi, C., & Pickering, M. J. (2014). Prediction at all levels: Forward model predictions can enhance comprehension. *Language, Cognition and Neuroscience*, 29(1), 46-48. <https://doi.org/10.1080/01690965.2013.852229>
- Gómez, C. M., Marco, J., & Grau, C. (2003). Preparatory visuo-motor cortical network of the contingent negative variation estimated by current density. *Neuroimage*, 20, 216-224.
[https://doi.org/10.1016/S1053-8119\(03\)00295-7](https://doi.org/10.1016/S1053-8119(03)00295-7)
- Gosselke Berthelsen, S., Horne, M., Brännström, K. J., Shtyrov, Y., & Roll, M. (2018). Neural processing of morphosyntactic tonal cues in second-language learners. *Journal of Neurolinguistics*, 45, 60-78. <https://doi.org/10.1016/j.jneuroling.2017.09.001>
- Grisoni, L., Miller, T. M., & Pulvermuller, F. (2017). Neural correlates of semantic prediction and resolution in sentence processing. *Journal of Neuroscience*, 37(18), 4848-4858.
<https://doi.org/10.1523/JNEUROSCI.2800-16.2017>
- Grisoni, L., Tomasello, R., & Pulvermuller, F. (2021). Correlated brain indexes of semantic prediction and prediction error: Brain localization and category specificity. *Cerebral Cortex*, 31(3), 1553-1568. <https://doi.org/10.1093/cercor/bhaa308>
- Gwilliams, L., & Davis, M. H. (2022). Extracting language content from speech sounds: The information theoretic approach. In L. L. Holt, J. E. Peelle, A. B. Coffin, A. N. Popper, & R. R. Fay (Eds.), *Speech Perception* (pp. 113-139). Springer International Publishing. https://doi.org/10.1007/978-3-030-81542-4_5
- Hed, A., Schremm, A., Horne, M., & Roll, M. (2019). Neural correlates of second language acquisition of tone-grammar associations. *The Mental Lexicon*, 14(1), 98-123.
<https://doi.org/10.1075/ml.17018.hed>
- Heilbron, M., & Chait, M. (2018). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience*, 389, 54-73.
<https://doi.org/10.1016/j.neuroscience.2017.07.061>
- Hickok, G. (2012). The cortical organization of speech processing: Feedback control and predictive coding the context of a dual-stream model. *Journal of Communication Disorders*, 45(6), 393-402. <https://doi.org/10.1016/j.jcomdis.2012.06.004>
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67-99. <https://doi.org/10.1016/j.cognition.2003.10.011>

- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177-180.
<https://doi.org/10.1126/science.182.4108.177>
- Hjortdal, A., Frid, J., & Roll, M. (2022). Phonetic and phonological cues to prediction: Neurophysiology of Danish stød *Journal of Phonetics*, *94*, 101178.
<https://doi.org/10.1016/j.wocn.2022.101178>
- Holmberg, A., & Platzack, C. (1995). *The role of inflection in Scandinavian syntax*. Oxford University Press.
- Hsu, Y.-F., Hämäläinen, J. A., & Waszak, F. (2016). The auditory N1 suppression rebounds as prediction persists over time. *Neuropsychologia*, *84*, 198–204.
<https://doi.org/10.1016/j.neuropsychologia.2016.02.019>
- Hunter, C. R. (2013). Early effects of neighborhood density and phonotactic probability of spoken words on event-related potentials. *Brain and Language*, *127*, 463-474.
<https://doi.org/10.1016/j.bandl.2013.09.006>
- Hunter, C. R. (2016). Is the time course of lexical activation and competition in spoken word recognition affected by adult aging? An event-related potential (ERP) study. *Neuropsychologia*, *91*, 451-464.
<https://doi.org/10.1016/j.neuropsychologia.2016.09.007>
- Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229-254). John Benjamins. <https://doi.org/10.1075/tsl.45.13jur>
- Kellmeyer, P., Ziegler, W., Peschke, C., Eisenberger, J., Schnell, S., Baumgaertner, A., Weiller, C., & Saur, D. (2013). Fronto-parietal dorsal and ventral pathways in the context of different linguistic manipulations. *Brain and Language*, *127*, 241-250.
<https://doi.org/10.1016/j.bandl.2013.09.011>
- Khanna, A., Pascual-Leone, A., Michel, C. M., & Farzan, F. (2015). Microstates in resting-state EEG: current status and future directions. *Neuroscience and Biobehavioral Reviews*, *49*, 105-113. <https://doi.org/10.1016/j.neubiorev.2014.12.010>
- Kujala, A., Alho, K., Service, E., Ilmoniemi, R. J., & Connolly, J. F. (2004). Activation in the anterior left auditory cortex associated with phonological analysis of speech input: localization of the phonological mismatch negativity response with MEG. *Cognitive Brain Research*, *21*(106-113). <https://doi.org/10.1016/j.cogbrainres.2004.05.011>
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32-59.
<https://doi.org/10.1080/23273798.2015.1102299>
- Lange, K., & Röder, B. (2006). Orienting attention to points in time improves stimulus processing both within and across modalities. *Journal of Cognitive Neuroscience*, *18*(5), 715-729. <https://doi.org/10.1162/jocn.2006.18.5.715>

- Lange, K., Rösler, F., & Röder, B. (2003). Early processing stages are modulated when auditory stimuli are presented at an attended moment in time: An event-related potential study. *Psychophysiology*, *40*, 806-817. <https://doi.org/10.1111/1469-8986.00081>
- Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*, 609-621. [https://doi.org/10.1016/0013-4694\(80\)90419-8](https://doi.org/10.1016/0013-4694(80)90419-8)
- 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., 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>
- Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, *31*(1), 133-156. <https://doi.org/10.1080/03640210709336987>
- Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1-71. [https://doi.org/10.1016/0010-0277\(80\)90015-3](https://doi.org/10.1016/0010-0277(80)90015-3)
- McCarthy, G., & Donchin, E. (1976). The Effects of Temporal and Event Uncertainty in Determining the Waveforms of the Auditory Event Related Potential (ERP) [Article]. *Psychophysiology*, *13*(6), 581-590. <https://doi.org/10.1111/j.1469-8986.1976.tb00885.x>
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1-86. [https://doi.org/10.1016/0010-0285\(86\)90015-0](https://doi.org/10.1016/0010-0285(86)90015-0)
- Mento, G., Tarantino, V., Sarlo, M., & Bisiacchi, P. S. (2013). Automatic temporal expectancy: a high-density event-related potential study. *PLoS ONE*, *8*(5), e62896. <https://doi.org/10.1371/journal.pone.0062896>
- Mietz, A., Toepel, U., Ischebeck, A., & Alter, K. (2008). Inadequate and infrequent are not alike: ERPs to deviant prosodic patterns in spoken sentence comprehension. *Brain and Language*, *104*, 159-169. <https://doi.org/10.1016/j.bandl.2007.03.005>
- Myrberg, S. (2010). *The intonational phonology of Stockholm Swedish* [Dissertation, Stockholm University]. Stockholm.
- Nagai, Y., Critchley, H. D., Featherstone, E., Fenwick, P. B. C., Trimble, M. R., & Dolan, R. J. (2004). Brain activity relating to the contingent negative variation: an fMRI investigation. *Neuroimage*, *21*(4), 1232-1241. <https://doi.org/10.1016/j.neuroimage.2003.10.036>

- Newman, R. L., & Connolly, J. F. (2009). Electrophysiological markers of pre-lexical speech processing: Evidence for bottom-up and top-down effects on spoken word processing. *Biological Psychology*, *80*(1), 114-121.
<https://doi.org/https://doi.org/10.1016/j.biopsycho.2008.04.008>
- Newman, R. L., Connolly, J. F., Service, E., & McIvor, K. (2003). Influence of phonological expectations during a phoneme deletion task: evidence from event-related brain potentials. *Psychophysiology*, *40*(4), 640-647. <https://doi.org/10.1111/1469-8986.00065>
- Nobre, A. C., & van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nat Rev Neurosci*, *19*(1), 34-48. <https://doi.org/10.1038/nrn.2017.141>
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189-234. [https://doi.org/10.1016/0010-0277\(94\)90043-4](https://doi.org/10.1016/0010-0277(94)90043-4)
- Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357-395. <https://doi.org/10.1037/0033-295X.115.2.357>
- Norris, D., McQueen, J. M., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Language, Cognition and Neuroscience*, *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* [Doctoral thesis, Lund University]. Lund.
- Novén, M., Olsson, H., Helms, G., Horne, M., Nilsson, M., & Roll, M. (2021). Cortical and white matter correlates of language-learning aptitudes. *Human Brain Mapping*, *42*(15). <https://doi.org/10.1002/hbm.25598>
- Novén, M., Schremm, A., Horne, M., & Roll, M. (2021). Cortical thickness of left anterior temporal areas affect processing of phonological cues in native speakers. *Brain Research*, *1750*, 147150. <https://doi.org/10.1016/j.brainres.2020.147150>
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*(4), 375-425. <https://doi.org/10.1111/j.1469-8986.1987.tb00311.x>
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785-806.
[https://doi.org/10.1016/0749-596X\(92\)90039-Z](https://doi.org/10.1016/0749-596X(92)90039-Z)
- Peña, J. M. (2022). Stød timing and domain in Danish. *Languages*, *7*(1), 50.
<https://doi.org/10.3390/languages7010050>
- Pulvermüller, F., & Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *Neuroimage*, *20*, 159-172.
[https://doi.org/10.1016/s1053-8119\(03\)00261-1](https://doi.org/10.1016/s1053-8119(03)00261-1)

- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79-87. <https://doi.org/10.1038/4580>
- Riad, T. (2014). *The phonology of Swedish*. Oxford University Press.
- Rischel, J. (1963). Morphemic tone and word tone in Eastern Norwegian. *Phonetica*, 10, 154-164. <https://doi.org/10.1159/000258166>
- Rocchetta, A. I. d., & Milner, B. (1993). Strategic search and retrieval inhibition: The role of the frontal lobes. *Neuropsychologia*, 31(6), 503-524. [https://doi.org/10.1016/0028-3932\(93\)90049-6](https://doi.org/10.1016/0028-3932(93)90049-6)
- Rodriguez-Fornells, A., Clahsen, H., Lleó, C., Zaake, W., & Münte, T. F. (2001). Event-related brain responses to morphological violations in Catalan. *Cognitive Brain Research*, 11(1), 47-58. [https://doi.org/https://doi.org/10.1016/S0926-6410\(00\)00063-X](https://doi.org/https://doi.org/10.1016/S0926-6410(00)00063-X)
- Roll, M. (2006). Prosodic cues to the syntactic structure of subordinate clauses in Swedish. In G. Bruce & M. Horne (Eds.), *Nordic Prosody IX* (pp. 295-204). Peter Lang.
- Roll, M. (2015). A neurolinguistic study of South Swedish word accents: Electrical brain potentials in nouns and verbs. *Nordic Journal of Linguistics*, 38, 149-162. <https://doi.org/10.1017/S0332586515000189>
- Roll, M. (2022). The predictive function of Swedish word accents. *Frontiers in Psychology*, 13, 910787. <https://doi.org/10.3389/fpsyg.2022.910787>
- Roll, M., & Horne, M. (2011). Interaction of right- and left-edge prosodic boundaries in syntactic parsing. *Brain Research*, 1402, 93-100. <https://doi.org/10.1016/j.brainres.2011.06.002>
- Roll, M., Horne, M., & Lindgren, M. (2009). Left-edge boundary tone and main clause verb effects on syntactic processing in embedded clauses—An ERP study. *Journal of Neurolinguistics*, 22(1), 55-73. <https://doi.org/10.1016/j.jneuroling.2008.06.001>
- Roll, M., Horne, M., & Lindgren, M. (2010). Word accents and morphology—ERPs of Swedish word processing. *Brain Research*, 1330, 114-123. <https://doi.org/10.1016/j.brainres.2010.03.020>
- Roll, M., Horne, M., & Lindgren, M. (2011). Activating without inhibiting: Left-edge boundary tones and syntactic processing. *Journal of Cognitive Neuroscience*, 23(5), 1170-1179. <https://doi.org/10.1162/jocn.2010.21430>
- Roll, M., Söderström, P., Frid, J., Mannfolk, P., & Horne, M. (2017). Forehearing words: Pre-activation of word endings at word onset. *Neuroscience Letters*, 658, 57-61. <https://doi.org/10.1016/j.neulet.2017.08.030>
- Roll, M., Söderström, P., & Horne, M. (2013). Word-stem tones cue suffixes in the brain. *Brain Research*, 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 and Language*, 150, 14-21. <https://doi.org/10.1016/j.bandl.2015.07.009>
- Sanders, L. D., & Astheimer, L. B. (2008). Temporally selective attention modulates early perceptual processing: Event-related potential evidence. *Perception and Psychophysics*, 70, 732–742. <https://doi.org/10.3758/PP.70.4.732>
- Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, 15, 228–240. [https://doi.org/10.1016/S0926-6410\(02\)00195-7](https://doi.org/10.1016/S0926-6410(02)00195-7)
- Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language*, 137, 29-39. <https://doi.org/10.1016/j.bandl.2014.07.010>
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(46), 18035-18040. <https://doi.org/10.1073/pnas.0805234105>
- Schafer, E. W. P., & Marcus, M. M. (1973). Self-Stimulation Alters Human Sensory Brain Responses. *Science*, 181(4095), 175. <https://doi.org/10.1126/science.181.4095.175>
- Schremm, A., Novén, M., Horne, M., Söderström, P., Westén, D. v., & Roll, M. (2018). Cortical thickness of planum temporale and pars opercularis in native language tone processing. *Brain and Language*, 176, 42-47. <https://doi.org/10.1016/j.bandl.2017.12.001>
- Schröger, E., Marzecová, A., & SanMiguel, I. (2015). Attention and prediction in human audition: A lesson from cognitive psychophysiology. *European Journal of Neuroscience*, 41, 641–664. <https://doi.org/10.1111/ejn.12816>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Shtyrov, Y., Pihko, E., & Pulvermüller, F. (2005). Determinants of dominance: Is language laterality explained by physical or linguistic features of speech? *Neuroimage*, 27(1), 37-47. <https://doi.org/10.1016/j.neuroimage.2005.02.003>
- Skeide, M. A., Brauer, J., & Friederici, A. D. (2016). Brain functional and structural predictors of language performance. *Cerebral Cortex*, 26(5), 2127-2139. <https://doi.org/10.1093/cercor/bhv042>
- Skeide, M. A., & Friederici, A. D. (2016). The ontogeny of the cortical language network [Perspective]. *Nature Reviews Neuroscience*, 17, 323. <http://dx.doi.org/10.1038/nrn.2016.23>

- Sumner, P., Nachev, P., Morris, P., Peters, A. M., Jackson, S. R., Kennard, C., & Husain, M. (2007). Human medial frontal cortex mediates unconscious inhibition of voluntary action. *Neuron*, 54(5), 697-711. <https://doi.org/10.1016/j.neuron.2007.05.016>
- Söderström, P., Horne, M., Frid, J., & Roll, M. (2016). Pre-activation negativity (PrAN) in brain potentials to unfolding words. *Frontiers in Human Neuroscience*, 10, 1-11. <https://doi.org/10.3389/fnhum.2016.00512>
- Söderström, P., Horne, M., Mannfolk, P., Westen, D. v., & 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 and Language*, 174, 119-126. <https://doi.org/10.1016/j.bandl.2017.08.004>
- Söderström, P., Horne, M., Mannfolk, P., Westen, D. v., & Roll, M. (2018). Rapid syntactic pre-activation in Broca's area: Concurrent electrophysiological and haemodynamic recordings. *Brain Research*, 1697, 76-82. <https://doi.org/10.1016/j.brainres.2018.06.004>
- Söderström, P., Horne, M., & Roll, M. (2016). Word accents and phonological neighbourhood as predictive cues in spoken language comprehension. *Speech Prosody 2016*, Boston.
- Söderström, P., Horne, M., & Roll, M. (2017). Stem tones pre-activate suffixes in the brain. *Journal of Psycholinguistic Research*, 46, 271-280. <https://doi.org/10.1007/s10936-016-9434-2>
- Söderström, P., Roll, M., & Horne, M. (2012). Processing morphologically conditioned word accents. *The Mental Lexicon*, 7(1), 77-89. <https://doi.org/10.1075/ML.7.1.04SOE>
- van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience*, 13(7), 967-985. <https://doi.org/10.1162/089892901753165872>
- Wacongne, C., Changeux, J.-P., & Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *The Journal of Neuroscience*, 32(11), 3665-3678. <https://doi.org/10.1523/jneurosci.5003-11.2012>
- Wacongne, C., Labyt, E., Wassenhove, V. v., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences*, 108(51), 20754-20759. <https://doi.org/10.1073/pnas.1117807108>
- Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensory-motor association and expectancy in the human brain. *Nature*, 203, 380-384. <https://doi.org/10.1038/203380A0>
- Yildiz, I. B., von Kriegstein, K., & Kiebel, S. J. (2013). From birdsong to human speech recognition: Bayesian inference on a hierarchy of nonlinear dynamical systems. *PLoS*

Computational Biology, 9(9), e1003219.

<https://doi.org/10.1371/journal.pcbi.1003219>