Good-enough parsing, activation, prediction: Evidence from EEG

Philippe Blache, Stéphane Rauzy, Sophie Dufour, Chotiga Pattamadilok, Deirdre Bolger

Good-enough parsing, activation, prediction: Evidence from EEG

Philippe Blache, Stéphane Rauzy, Sophie Dufour, Chotiga Pattamadilok, Deirdre Bolger

Laboratoire de Parole et Langage (LPL)
Brain & Language Research Institute (BLRI)

Abstract

We present a parsing model called *good enough parsing* that integrates cue-based prediction and shallow processing techniques, based on a smooth matching of the predicted elements. We describe an EEG experiment showing that brain activity confirms the predictions of the model.

Background

The role of prediction is central to parsing and different techniques can be used for its implementation, such as lookahead, left-corner or transitional probabilities (Abney & Johnson, 1991; Resnik, 1992; van Schijndel et al., 2013). Prediction by anticipating at each time step the next word given a prior context reduces the search space and controls the number of possible solutions.

The prediction process makes use of different *cues* from such sources as syntax, pragmatics, prosody, context, gestures etc. Thus, prediction can activate words as well as sets of words, including over a long distance, thus implementing the notion of ***flexible incrementality*** (Reitter et al, 2011) in which processing is not strictly incremental (word-by-word) and interpretation is not always compositional.

Studies have shown that processing can become shallow in highly predictive contexts (Levy, 2008a, Rommers et al., 2013); these observations form the basis of the ***Good-enough Theory*** (Ferreira & Patson, 2007).

Prediction and Activation

Classical parsing uses forward transitional probabilities: the capacity to predict a word depends on its previous context. Several possibilities need to be integrated into the on-going syntactic structure and, thus, different expectances are generated (Hale, 2001). Such a word-by-word mechanism is usually based on morpho-syntactic information.

Prediction also plays a role in cognitive approaches. In memory-based architectures such as ART-R, several ***chunks*** are stored in working memory and, depending on their ***level of activation*** can be retrieved and integrated into structure. In our approach, cues can be of any type (syntactic, prosodic, pragmatic etc.) and not linked to the strict prior context. Chunks can represent any type of objects, from atomic to complex structures (i.e. lexical entries as well as entire constructions).

Cue-based Activation, Good-enough parsing

Here we propose to generalise predictive elements to any type of information.

- Any type of cue can be used
- Cues can interact and form a set of associated cues.
- The prediction can be triggered at any time, when the predictive information reaches a certain threshold.

So what precisely are these cues and how can they activate a construction (a *form-meaning* pair) ? We tackle this question by looking at the example of **idiomatic constructions**, which can be described as a set of words with a **generally fixed order** and with **little or no variability**. Such information can be described using a set of properties (Blache et al., 2006):

- *linear order between forms*
- *their possible adjacency,*
- *the possibility of introducing a modifier,*
- *the possibility of morphological variation etc.*

Example: In the following idiom, some variation is possible; the facultative elements are indicated in brackets:

"{don't} put all {your} eggs in [one | the same] basket".

This idiom can be represented by the following structure that encodes **the set of forms**, **their order**, the fact that some forms are **adjacent**, the idiom **meaning**:

$$\left[\begin{array}{ll} \text{FORM} & \left[\begin{array}{l} \text{MORPH}\ \{put,\ all,\ eggs,\ in,\ basket\} \\ \text{LIN}\ \{put \prec all \prec eggs \prec in \prec basket\} \\ \text{ADJ}\ \{put \oplus all;\ eggs \oplus in\} \end{array}\right] \\ \text{MEANING} & [\![S = \text{TAKE-A-RISK}]\!] \end{array}\right]$$

Each relation is weighted and constitutes a *cue* for the construction that can be integrated into the ***spreading activation*** term of the activation formula:

$$\sum_j W_j S_{ji}$$

in which a cue j (of weight W_j) is one of the relations described in the description of construction i. According to the above example the following cues are accessible upon reaching the form *« eggs »* :

{morph = [put, all, eggs]; put ≺ all; all ≺ eggs; put ⊕ all}.

Assigning a heavy weight to these cues, leads to a high activation strength that activates the idiomatic description and its relations.

In a **second stage of the parsing process**, the predicted structure is matched to the input. As the general organization of the construction (and its interpretation) is already available, we no longer speak of *parsing* but of ***smooth matching*** between the input word and the predicted form.

...	put	all	eggs	in	one	basket	...
scan →	•	•	•				

Predictions of this model regarding human sentence processing :

- A construction activator is a facilitator and leads to easier processing.
- This facilitator integrates a *compensation effect*: an unexpected difficulty is compensated by the prediction.
- Instead of recovering the error, *Good-enough parsing relaxes the violated relations and continues the processing.*

EEG Study

In our EEG study we analyse the brain activity in response to a syntactic violation introduced into idioms. We compare event-related potentials (ERP) of idioms and control sentence, with and without violations. Our model predicts that the difficulty of processing the violation will be compensated by the activation of an idiom : *the loose unification involved in the GE-parsing recovers the difficulty by relaxing the constraints.*

In the EEG experiment participants were presented with **120 French idioms** (ID), 60 with violations (IDV) and 60 without (IDNV), and **120 control sentences** (CTR), 60 with violations (CTRV) and 60 without (CTRNV). The distribution of idiom familiarity and violation type was controlled. The stimuli were presented word-by-word on-screen during EEG acquisition. ERPs were calculated for 3 positions :

1. **Recognition Point (RP) :** where the idiom is presumed to be recognised.
2. **Modified Word (MM)** : where the violation is introduced.
3. **Detection Word (MD1) :** where the violation is detected for the CTRV.

Results

Subject-level, trial-averaged EEG data was extracted for the three word positions for both Controls and Idioms with and without violations. A two-tailed cluster-based permutation test was carried out on to compare non-violation conditions (CTRNV and IDNV) and violation conditions (CTRV and IDV) for each word position (RP, MM and MD1).

Recognition Point (RP): As no effect of violation was expected at this position, the violation conditions were collapsed for both CTRL and ID conditions ((CTRNV+CTRV) vs. (IDNV+IDV)).

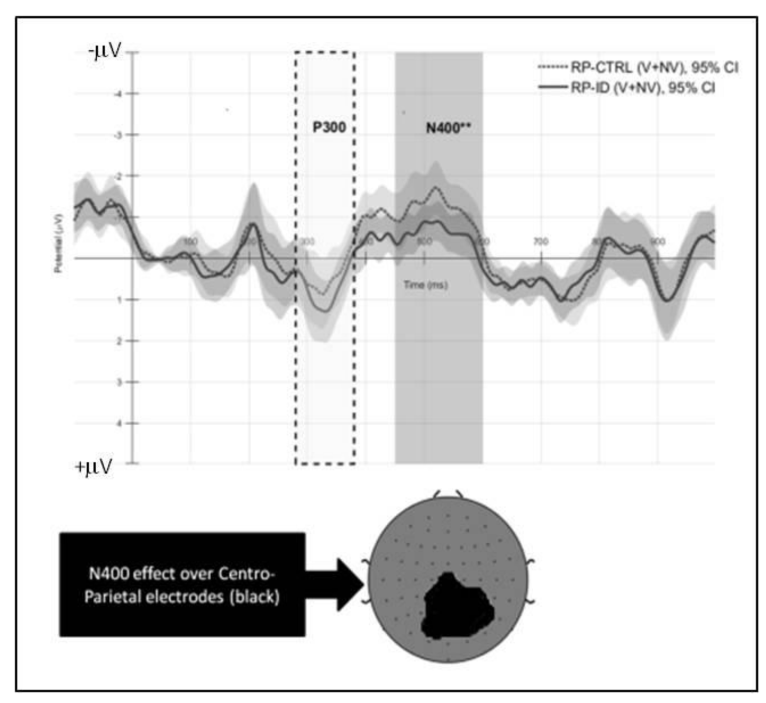

Figure 1: Position RP

Statistical analyses revealed a significant ($p \leq 0,025$) N400 difference over centro-parietal electrodes from 390 to 550ms; ctrl presented a higher N400 amplitude than ID (figure 1). This observatio is in line with previous findings of a reduced N400 in the context of idioms compared to literal sentences (Rommers et al., 2013) and is indicative of higher word-probability at RP for ID compared to CTRL.

A greater P300 effect, posited as an index of prediction processes in idioms (Molinaro & Carreiras, 2010) was observed for ID compared to CTRL. However, this did not reach statistical significance according to the cluster-based permutation test (figure 1).

Modified Word (MM): The violation condition in CTRL and ID were analysed separately. As expected, no significant difference was revealed for CTRL (CTRNV vs CTRV). However for ID, IDV presented a significantly higher N400 ($p \leq 0,025$) than IDNV (figure 2).

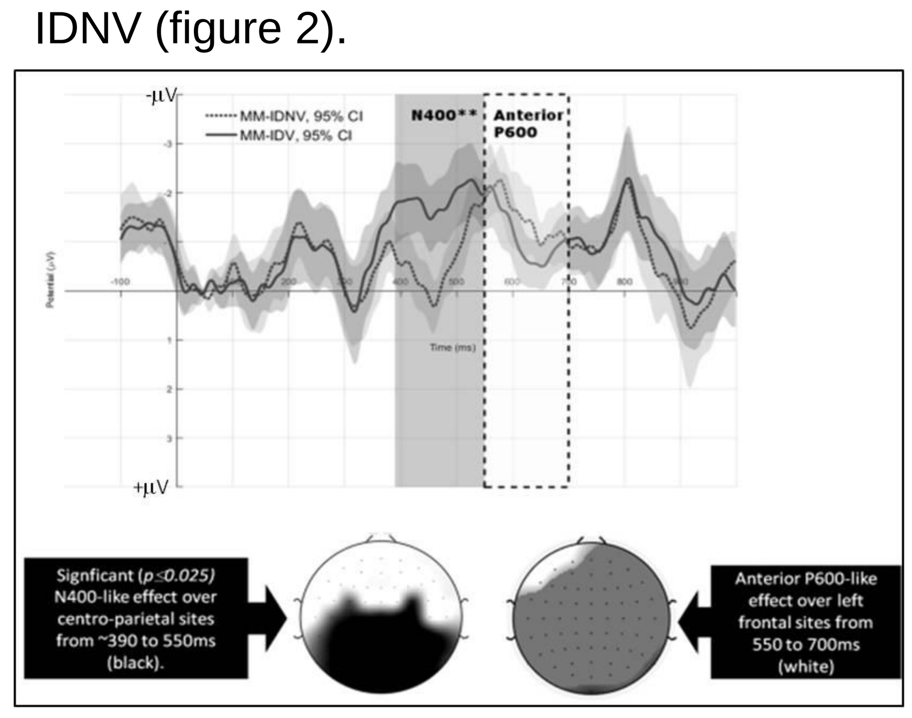

Figure 2: Position MM

A significant difference ($p \leq 0,025$) between IDV and IDNV in the 550 to 700ms window over left-frontal electrodes was also observed. IDV presented more positive-going activity compared to IDNV. This could be interpreted in light of the suggestion by Hagoort et al. (1999) that more frontally distributed P600-like effects may reflect an over-writing of an « *active structural representation* ».

Detection (MD1): At this position the reader detects the the violation introduced at position MM for CTRL. A CTRV vs. CTRNV comparison revealed a significant N400 effect ($p \leq 0,05$) (figure 3, left); this reflects the processing of the control violation.

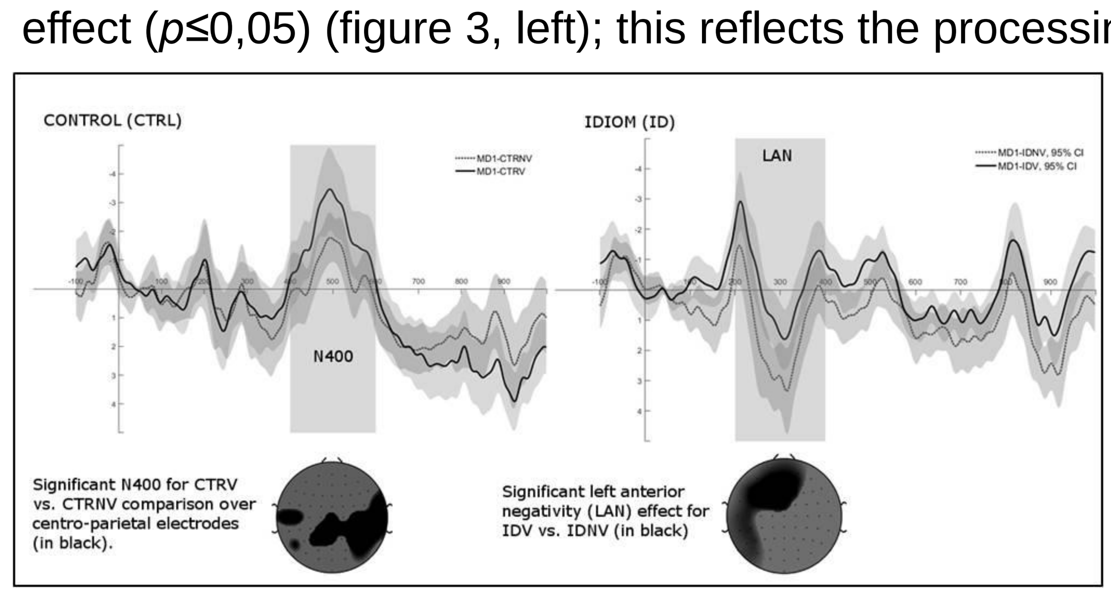

Figure 3: Position MD1

The N400 effect is followed by positivity for ctrv from around 600ms; this P600-like effect suggests the processing of the syntactic violation.

ID presents a reduced N400 for both IDNV and IDV and no significant N400 difference. However, there is evidence of a LAN (Left Anterior Negativity) over left frontal electrodes from 200ms to 400ms.

Conclusions

The EEG results validate the hypothesis of ***GE-parsing*** :

- Higher positivity (reduced N400, higher P300) for ID from RP validates the facilitator effect after RP due to the prediction of entire construction.
- The violation in ID (IDV) is detected at MM as indexed by significantly reduced N400.
- The violation in ID is recovered as suggested by frontal positivity in P600 time window.
- For CTRL, the violation is processed at MD1 as indexed by significant N400 for CTRV.
- For ID, the LAN at MDI suggests that the violation was detected but this violation is not repaired, as suggested by the lack of later positivity (P600).