Comprehenders generate long-distance predictions during reading: ERP evidence from verb-particle constructions

Kate Stone, Shravan Vasishth, Titus von der Malsburg (University of Potsdam)
stone@uni-potsdam.de

ERP evidence suggests that features of upcoming words are preactivated when the context is constraining [1-7, cf. 8]. It is unclear, however, under which circumstances this preactivation leads to a specific lexical prediction, i.e. the selection of just one lexical entry which is then integrated into the sentence interpretation ahead of time. Particle verb constructions are ideally suited for investigating this question [9]. In German particle verbs, a base verb occurs in second position (V2 word order) whereas a particle that gives the verb its full meaning remains in sentence-final position (see example on page 2). Unlike in English, there can be an almost arbitrary amount of material between the base verb and the particle. Lexical prediction of the particle can therefore facilitate incremental sentence interpretation by giving early access to verb semantics.

To investigate whether such long-distance lexical predictions are being made, we used sentences that either constrained the set of plausible particles to just one (condition a) or to a small set of two or more highly plausible particles of which just one was shown (c). In addition, we used two violation conditions (b, d) where the presented particle was not compatible with the context. We hypothesized that the parser would more likely commit to a lexical prediction when there was only one high-cloze particle (low entropy, a and b) than when there were multiple high-cloze particles (high entropy, c and d). Any ERP differences in (a) vs. (c) could be due to differences in cloze probability. We therefore focused on the violation conditions (particle cloze = 0) and predicted that a violation would cause greater difficulty (indexed by an N400 effect) in the presence of a specific lexical prediction (b), than when a commitment presumably had not been made (d). The study was pre-registered on OSF: https://osf.io/m96cq/

Methods: EEG was recorded from 50 participants with 44 target items and 62 filler sentences. RSVP was used at 190 ms/word + 20 ms/letter and 700 ms for the target particle, with an ISI of 300 ms. A norming study showed that cloze probability of particles in (a) was 90% and the difference between the top two particles (when two were given) was 74%. In (c), cloze probability was 73% and the top-two difference 29%. Pre-registered analysis: ERPs were analyzed using Bayesian LMMs that modeled by-trial mean amplitude. A sanity check established that violations elicited the expected N400 and late positivity (b/d vs a/c). The N400 prediction was tested at electrode Pz in the window 250-500 ms. Exploratory analysis: To test for late positivities which follow N400 effects in some studies of context-based word predictability [10], an additional analysis was conducted in the window 600-900 ms at electrode Cz. Deviations from pre-registration: 10 additional subjects were recruited; no data were analysed prior to extending recruitment. Bayes factors were not computed due to vague priors.

Results: Contrary to our prediction, there was no indication of an N400 difference between the two violation conditions (b vs d, $\hat{\beta} = -0.25\mu V$, $Crl = [-1.21, 0.72]\mu V$, $Pr(\beta < 0) = 0.71$, see Fig. 1). This result on its own could be interpreted as suggesting no lexical prediction of distant verb particles. However, as shown in Fig. 2, we found an effect in the time window 600-900 ms: a more positive-going amplitude in the violation condition with just one plausible particle (b) than in the violation condition with two or more plausible particles (d) ($\hat{\beta} = 0.96\mu V$, $Crl = [-0.20, 2.11]\mu V$, $Pr(\beta > 0) = 0.95$, see also Fig. 3).

Conclusions: Late positive components have been observed for anomalous words in strongly constraining contexts [10]. They are thought to reflect attempts at reanalysis, revision, and repair, depending on their topography [7,11]. We propose that in the 1-particle condition (b), a lexical prediction was triggered and a richer mental representation of the sentence built before the particle was seen. This representation then had to be revised or discarded once the violating particle was encountered and the late positivity reflects this cost. To establish the reliability of our findings, we are currently preparing a pre-registered replication with high statistical power. Tentatively, we conclude that German native speakers make long-distance lexical predictions if constraint is not just high but also strongly favors a single lexical item.
a. 1-particle/grammatical:

Die gemeine Bande hing dem unschuldigen Mann eine schreckliche Straftat völlig skrupellos an, einfach um Unruhe zu stiften.

b. 1-particle/ungrammatical:

*Die gemeine Bande hing dem unschuldigen Mann eine schreckliche Straftat völlig skrupellos nieder, einfach um Unruhe zu stiften.

*The nasty gang blamed the innocent man of a terrible crime completely without scruple at/*down, simply to cause trouble.

c. 2+particle/grammatical:

Die gemeine Tante hing das schreckliche Porträt von ihrem Mann völlig skrupellos ab/auf, einfach um Unruhe zu stiften.

d. 2+particle/ungrammatical:

*Die gemeine Tante hing das schreckliche Porträt von ihrem Mann völlig skrupellos nieder, einfach um Unruhe zu stiften.

*The mean aunt hung the terrible portrait of her husband completely without scruple off/up /*down, simply to cause trouble.

Figure 1: ERPs elicited by violation particles. Shaded is the pre-registered N400 analysis window.

Figure 2: Shaded is the analysis window of the late positivity.

Figure 3: Topographical plot of condition (b) minus condition (d), 600-900 ms.