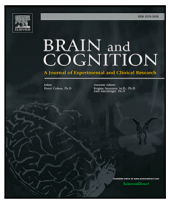
Brain and Cognition 186 (2025) 106293

Contents lists available at ScienceDirect

Brain and Cognition

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



Full length article

# On the biphasic nature of the N400-P600 complex underlying language comprehension

Francesca Delogu [a,*], Christoph Aurnhammer [a], Harm Brouwer [a,b], Matthew W. Crocker [a]

[a] Department of Language Science and Technology, Saarland University, Saarbrücken, D-66123, Germany
[b] Department of Cognitive Science and Artificial Intelligence, Tilburg University, Tilburg, The Netherlands

ARTICLE INFO

Dataset link: https://osf.io/wgjxh/

*Keywords:*
Language comprehension
Event-Related Potentials (ERPs)
N400
P600
Task
Spatiotemporal component overlap

ABSTRACT

The ERP literature on language comprehension reveals variability in observing monophasic N400 versus biphasic N400-P600 effects in response to incongruent input, with the reasons for this inconsistency remaining unclear. Two interrelated factors may contribute: spatiotemporal overlap between the N400 and P600, where a strong N400-effect can obscure the P600, and the P600's sensitivity to depth of processing, as determined by the experimental setting. Building on previous findings reporting monophasic N400-effects with plausibility judgments, we investigated whether comprehension questions, encouraging more natural reading and deeper processing of the full content, would elicit a biphasic effect, suggesting reduced component overlap in such settings. Using a design fully crossing lexical association and plausibility, we found that the N400 is modulated by association and the P600 by plausibility. Crucially, a biphasic pattern emerged for implausible and unrelated words, suggesting a mitigation of component overlap compared to previous studies employing plausibility judgments. We interpret the results in light of current accounts of the N400 and P600, arguing that the empirical evidence strongly supports single-stream over multi-stream models. Importantly, our findings highlight the critical role of both component overlap and task demands in shaping the data that inform the development and evaluation of theoretical models.

## 1. Introduction

In ERP research on language comprehension, semantic violation paradigms are often employed to explore how word meaning is retrieved and integrated into broader linguistic contexts. Typically, semantically incongruent words elicit an N400 effect relative to congruent words, a response that has been associated with processes related to semantic integration (Brown & Hagoort, 1993, 1999), lexical retrieval/access (Kutas & Federmeier, 2000, 2011), or a combination of both (Baggio & Hagoort, 2011; Nieuwland et al., 2019). Conversely, several studies investigating anomalous sentences such as those including role reversal violations have reported a P600 effect without an N400 effect – a phenomenon known as the 'semantic P600' (see Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer et al., 2012; Kuperberg et al., 2007, for reviews). The full spectrum of extant data has motivated two main classes of accounts. Multi-stream models posit that the N400 and P600 reflect two distinct processing streams – heuristic and algorithmic, respectively – that build interpretations in parallel (see Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer et al., 2012; Kuperberg, 2007, for reviews). On these models, the N400 reflects difficulty in constructing a structure-independent heuristic semantic interpretation of a sentence, while the P600 is evoked whenever this semantic interpretation mismatches with that of the structure-dependent algorithmic one. As both processing streams should agree upon the implausibility of semantic incongruities, multi-stream models typically predict only a monophasic N400 for semantic incongruities. By contrast, single-stream models, such as the Retrieval-Integration (RI) account (Brouwer et al., 2017, 2012), posit that the N400 indexes lexical retrieval/access, and the P600 reflects the integration of retrieved word meaning into the unfolding utterance representation.

While multi-stream models – by design – predict that traditional semantic incongruities result in a monophasic N400 effect (see Brouwer et al., 2012, for a comprehensive discussion of these architectures), single-stream accounts such as the RI model predict these incongruities to elicit a biphasic N400-P600 pattern reflecting retrieval (N400) and integration (P600) effort, where integration effort is taken to reflect syntactic (e.g. Gouvea et al., 2010), semantic (e.g. Aurnhammer et al., 2023; Delogu et al., 2019, 2021), and pragmatic (e.g. Burkhardt, 2007; Hoeks & Brouwer, 2014; Regel et al., 2011) factors affecting the unfolding interpretation. Crucially, however, while the N400 effect does

* Corresponding author.
*E-mail address:* delogu@lst.uni-saarland.de (F. Delogu).

https://doi.org/10.1016/j.bandc.2025.106293
Received 18 December 2024; Received in revised form 13 March 2025; Accepted 13 March 2025
Available online 27 March 2025
0278-2626/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

indeed sometimes co-occur with a P600 effect, there is considerable variability in the presence of this positive component. For instance, a systematic review on semantic incongruities (Van Petten & Luka, 2012) found that only about one third of the studies they reviewed reported a biphasic N400-P600 effect, while the remainder observed only an N400 effect. This variability persists despite the studies being homogeneous with regard to factors such as the absence of an overt task, the syntactic well-formedness of the sentences, and the sentence-final position of the target word (see also Kim et al., 2024).

In order to reliably interpret ERP effects, it is essential that we understand those factors that underlie this variability. Van Petten and Luka (2012) suggest that the post-N400 positivity may reflect a reanalysis process that can be variably triggered by several factors, including the extent to which the incongruent sentence can be reinterpreted to make sense and the individual verbal skills or motivation of the participants (p. 184). Alternatively, Brouwer and Crocker (2017) propose that this variability may arise from the degree to which the P600 effect survives spatiotemporal overlap with the N400 component, since incongruent words typically also elicit increased N400 amplitudes. That is, scalp-recorded ERP waveforms represent the summation of various latent ERP components reflecting simultaneously occurring cognitive processes (Luck, 2005). When latent components have opposite polarities, like the N400 and the P600, and the processes underlying these components (partially) overlap in time, they may interfere with each other in the observed surface signal. RI theory predicts that the retrieval processes underlying the N400 may at least partially temporally overlap with the integration processes underlying the P600, leading to spatiotemporal component overlap of these components in the scalp-recorded ERP signal (Brouwer & Hoeks, 2013, p. 9). Indeed, Brouwer and Crocker (2017) argue that such spatiotemporal component overlap may lead to the apparent absence (or presence) of an effect in the observed waveforms. For example, the presence of an N400 elicited by semantically incongruent words, as in the studies reviewed by Van Petten and Luka (2012), may attenuate a later positivity, depending on the relative onset, amplitude, and duration of the two components.

Supporting the component overlap hypothesis, a series of studies investigating ERP effects of lexical association (retrieval) and plausibility (integration) have shown that whether or not a P600 effect for implausible targets is observed may depend on amplitude differences in the N400 time window: Whereas targets that are highly associated to the context yield no N400 effect but rather a P600 effect for plausibility, this P600 effect is absent for targets that are strongly unassociated, which elicit only an N400 effect in the observed waveforms (Delogu et al., 2019, henceforth DBC19). Critically, the P600 effect for implausibility is observed when targets are equally unassociated, thereby increasing N400 amplitude to a similar degree (Delogu et al., 2021, henceforth DBC21, see also Brouwer, Delogu, and Crocker, 2021). To mitigate the influence of component overlap on the P600, one approach could involve reducing the amplitude of the N400, for instance, by priming the target word expected to produce the P600 effect (Aurnhammer et al., 2023, DBC19). Another strategy, as Brouwer and Crocker (2017) hypothesize, would be to enhance the P600 by incorporating a task that induces deeper sentence comprehension, which has been shown to generally yield more pronounced P600 responses (e.g., Kolk et al., 2003; Schacht et al., 2014).

The present study addresses this hypothesis by examining whether the component overlap influence on the P600 – i.e., the attenuation of the P600 by a large N400 effect in DBC19 – persists with a more fully engaging and ecological task. While we do not directly manipulate the task in the present experiment, we assess whether using comprehension questions on a proportion of trials — encouraging participants to broadly recover sentence meaning rather than focus on a single aspect of it — results in a biphasic N400-P600 effect. Such an outcome would suggest a mitigation of component overlap in the observed signal compared to when plausibility judgments were used, as in DBC19.

To fully test these predictions, we adapted the designs employed in DBC19 and DBC21 into a 2×2 design fully crossing Association (Related vs. Unrelated) and Plausibility (Plausible vs. Implausible) of target words given the context, as can be seen in Table 1, while also replacing the simple judgment task with more engaging comprehension questions. This design enables us to assess the presence of a P600 effect of implausibility in both related and unrelated targets compared to their plausible controls, as well as to examine any potential interaction between association and plausibility. Additionally, the comparison between Unrelated-Implausible and Related-Plausible targets allows us to evaluate the degree of component overlap – that is, whether the P600 effect driven by implausibility is obscured or rather survives overlap with the N400 effect driven by association.

We expect to replicate previous findings, namely, an N400 effect for unrelated targets (*house door*) compared to related ones (e.g., *menu*), with no influence from plausibility, and a P600 effect for less plausible conditions compared to their plausible counterparts. We may also expect an interaction between association and plausibility, with the strength of the P600 effect (integration) varying as a function of the N400 (retrieval). Finally, we examine whether a task that requires participants to fully attend to the entire content of what has been read leads to a mitigation of the component overlap influence on the P600 in the observed signal, i.e., elicits a biphasic N400-P600 effect for Unrelated-Implausible targets relative to their Related-Plausible controls, which resulted in an N400 effect only in DBC19. Taken in the context of previous findings, these results would support single-stream accounts of language processing (Brouwer et al., 2017, 2012) and challenge the aforementioned alternative multi-stream models, which are typically unable to accommodate biphasic effects.

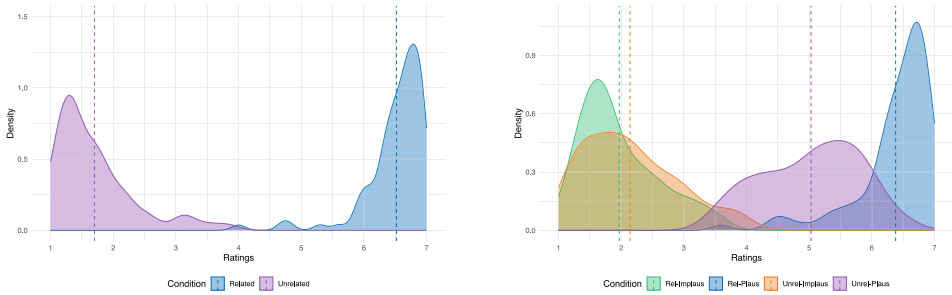
## 2. Material and methods

### 2.1. Participants

We report data from 28 participants from Saarland University. All were right-handed, native speakers of German, and had a normal or corrected-to-normal vision. Eight additional participants were excluded from the analysis due to excessive eye-movements or other artifacts. All participants gave written informed consent and were paid for taking part in the experiment. The reported studies were conducted in accordance with the ethics approval granted by the Deutsche Gesellschaft für Sprachwissenschaft (DGfS).

### 2.2. Materials

We constructed 100 German sentence-pairs in four conditions, as illustrated in Table 1 (see Appendix A for the original versions in German). Note that the materials used in this study were not exactly the same as those in DBC19 and DBC21. This is because the context sentences in those previous experiments were not identical, and therefore, it was not possible to simply combine the items to create the four conditions of the present study. Additionally, the added target manipulation made it necessary to control for target word length and frequency, which was not required in the earlier studies. The average length in number of characters of the target words was 7.42 (SD = 2.87) for the Related conditions, and 7.6 (SD = 3.01) for the Unrelated conditions. Word frequencies were extracted from the deWaC corpus of German (Baroni & Kilgarriff, 2006). No significant differences across conditions were observed (*ps* >.15). Apart from these changes, every effort was made to keep the stimuli as close as possible to those in DBC19 and DBC21. The materials in the present study were re-normed for semantic association, plausibility and cloze probability, showing similar patterns to those of the previous studies. The results of the norming studies are reported in Table 2. All participants in the pre-tests were recruited via Prolific Academic (www.prolific.co).




(a) Association ratings (b) Plausibility ratings

**Fig. 1.** Density plots showing the distribution of the association and plausibility ratings by condition, as collected in the norming studies.

**Table 1**
Example of the experimental materials (with English translation). The target sentence varied only between the Related and the Unrelated conditions. The target word is underlined for illustrative purposes only.

| Context sentence | Target sentence | Condition |
|---|---|---|
| Markus setzte sich im Restaurant.<br>*(Markus sat down in the restaurant.)* | Bald darauf öffnete er die Speisekarte und …<br>*(Before long, he opened the menu and …)* | Related-Plausible |
| Markus verließ das Restaurant.<br>*(Markus left the restaurant.)* | | Related-Implausible |
| Markus verließ das Restaurant.<br>*(Markus left the restaurant.)* | Bald darauf öffnete er die Wohnungstür und …<br>*(Before long, he opened the house door and …)* | Unrelated-Plausible |
| Markus setzte sich im Restaurant.<br>*(Markus sat down in the restaurant.)* | | Unrelated-Implausible |

**Table 2**
Results of the three norming studies. Mean scores are reported by condition, with standard deviations in parentheses.

| Condition | Related-Plausible | Related-Implausible | Unrelated-Plausible | Unrelated-Implausible |
|---|---|---|---|---|
| Cloze Probability | 0.50 (0.23) | 0.04 (0.06) | 0.004 (0.02) | 0.00 (0.00) |
| Association Rating | 6.52 (0.49) | 6.52 (0.49) | 1.70 (0.63) | 1.70 (0.63) |
| Plausibility Rating | 6.38 (0.62) | 1.97 (0.62) | 5.03 (0.78) | 2.14 (0.75) |

To estimate the cloze probability of the target words, four lists counterbalancing items and conditions were created with the sentence pairs presented up to and excluding the determiner preceding the target noun (e.g., "Markus setzte sich im Restaurant. Bald darauf öffnete er … "). Each list was assigned to 15 participants. The mean cloze probability of the target was higher in the Related-Plausible condition (.5) compared to the other conditions, where cloze probabilities were close to 0.

The semantic association between the nouns in the context (e.g., *Restaurant*) and the related vs. unrelated nouns in the target sentence (*Speisekarte* vs. *Wohnungstür*) was rated on a 7-points Likert scale (1 = not at all related, 7 = strongly related) by 60 participants. As expected, the related pairs were rated as more associated (6.5) than the unrelated ones (1.7).

Plausibility ratings were collected on a 7-points Likert scale (1 = highly implausible, 7 = highly plausible) for all the experimental items presented up to the target noun, to avoid judgments to be affected by sentential materials appearing after it. Four counterbalanced lists were created, each one presented to 10 participants who did not take part in the cloze or the association rating studies. The Related-Implausible condition was judged to be less plausible than the Related-Plausible condition, and the Unrelated-Implausible condition was rated as less plausible than the Unrelated-Plausible condition. The distributions of plausibility and association ratings by conditions are displayed in Fig. 1.

Four counterbalanced lists were created so that each item appeared in each list in a different condition. The experimental items were intermixed with 100 filler passages created to vary the verb in the context sentence, the type and locus of the implausibility (half of the fillers were implausible), and the sentence structure. Comprehension questions requiring a yes-no answer were created for one third of all items. Half of the questions addressed the context sentence and the other half the target sentence. Half of the questions required a "yes" answer, the other half a "no" answer. This design ensured that the questions were unpredictable both in terms of their content and their timing within the experiment. Examples of the questions are reported in Appendix A.

### 2.3. Procedure

Participants were seated in a dimly lit sound-proof, electro-magneti cally shielded booth, in front of a 24 inch computer screen. Stimuli were presented with the E-prime software (Psychology Software Tools, Inc.) in white font on a black background. Each trial began with a screen instructing participants to press a button to start reading. The context sentence was displayed as a whole until participants pressed a button to proceed. A fixation cross then appeared for 750 ms, followed by the target sentence presented word-by-word in the center of the screen. Each word was shown for 350 ms preceded by a 150 ms inter-stimulus interval. Next, either a comprehension question appeared, to which participants responded by pressing one of two buttons on a response box, or the next trial began. The position of the yes-no buttons on the left or right of the screen was randomized across trials. The experiment was divided in four blocks, with breaks between blocks. At the end

of each block, participants received feedback on their performance on the comprehension questions, in order to keep them motivated and engaged with the task.

### 2.4. Electrophysiological recording and processing

The EEG was recorded using 26 active scalp electrodes placed according to the 10–20 system. The horizontal electro-oculogram (EOG) was monitored with two electrodes placed at the outer canthi of each eye and the vertical EOG with two electrodes above and below the left eye. Electrode impedance was kept below 5 k$\Omega$ for all scalp electrode sites, and below 10 k$\Omega$ for the EOG electrodes. The signal was digitized at a sampling rate of 500 Hz. During recording no on-line filters were used.

### 2.5. Analyses

The EEG signal was band-pass filtered offline at 0.01–30 Hz and re-referenced to the average of the left and right mastoid electrodes. The EEG was segmented into epochs time-locked to the onset of the target nouns (−200 ms to 1200 ms). The baseline used a 200 ms pre-stimulus window. Trials contaminated by eye-movements, blinks, muscle activity, or other artifacts were removed prior to analysis. On average, there were 18 trials per condition following artifact rejection (range: 12–25), with no significant differences across conditions, F(3, 81) = 0.36, p = .78. We performed repeated-measures ANOVAs on mean amplitudes in the 300–500 ms (N400) and in the 600–1000 ms (P600) time windows for nine representative electrodes (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4). We also analyzed six consecutive 100 ms time windows starting at 400 ms, to assess how the effects develop over time. The results of these statistical analyses are shown in Appendix B.

The ANOVAs within each time window included Association (Related vs. Unrelated) and Plausibility (Plausible vs Implausible) as within-subject factors, along with two topographic factors reflecting the location of the nine electrodes on the scalp: Anterior-Posterior (AP) distribution (anterior, central, posterior) and Laterality (left, central, posterior). To directly assess component overlap, we performed planned comparisons between the Related-Plausible and the Unrelated-Implausible condition in both the N400 and the P600 time windows. The ANOVAs for this contrast included Condition, AP distribution, and Laterality as within-subject factors. The Greenhouse–Geisser correction was used in all ANOVAs with greater than one degree of freedom in the numerator. In such cases, the corrected $p$-value is reported.

We complemented these inferential analyses with a regression-based estimation technique (rERP) (Brouwer, Delogu, & Crocker, 2021; Smith & Kutas, 2015a, 2015b) to assess the relative contribution of plausibility and association to the scalp-recorded voltages on a trial-by-trial basis. We fitted multiple linear regression models for each participant, electrode, and 2 ms time-slice (corresponding to a 500 Hz sampling rate) separately, with association and plausibility ratings collected in the norming studies as continuous predictors. In order to aid interpretation of the slope coefficients, the plausibility and association predictors were inverted and z-standardized. As a result, higher association values indicate greater unassociation, and higher plausibility values correspond to increasing implausibility. After fitting the models on a trial-by-trial basis, trials were re-averaged by condition, and model fit was qualitatively assessed by inspecting mean residuals – calculated as the difference between the observed and the estimated data – per electrode and time-slice within each condition. We report the fitted coefficients, to visualize how the predictors combine in producing the scalp-recorded voltages, the associated effect size, and the residuals averaged by condition (see Brouwer, Delogu, & Crocker, 2021). As an indication of effect size, we report t-values computed from the same models across (rather than within) subjects. This allows us to obtain a single $t$-value for each electrode and time sample. The residuals – the difference between the observed data and the forward estimates computed by the model – will be inspected to assess how accurately the predictors model the trends in the data. The closer the residuals are to 0, the better the fit of the estimates to the observed data.

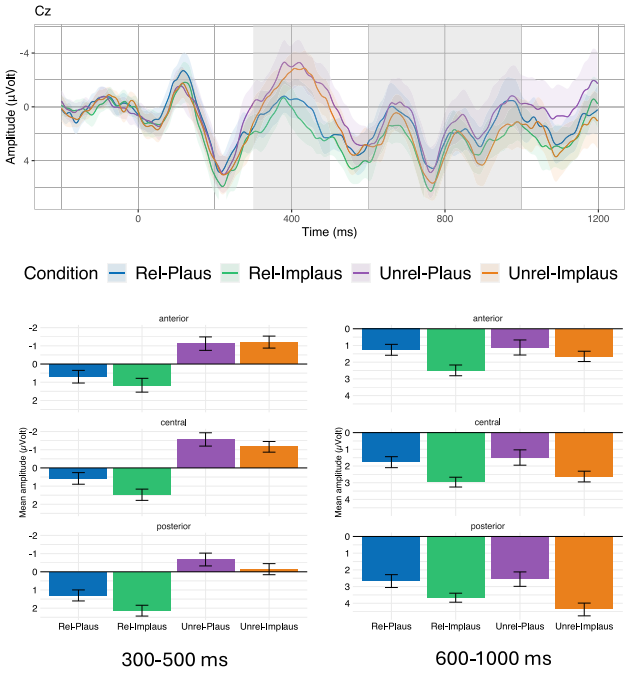



**Fig. 2. Top**: grand-average event-related potentials (ERPs) for the four conditions, plotted at Cz. In this and subsequent figures, negative is plotted up. **Bottom**: bar graphs showing average voltages on anterior (F3, Fz, F4), central (C3, Cz, C4) and posterior sites (P3, Pz, P4) within the N400 (300–500 ms) and the P600 (600–1000 ms) time windows, with ±1 standard error of the mean (SEM). In this and all subsequent figures, the conditions are labeled as follows: Rel-Plaus = Related-Plausible, Rel-Implaus = Related-Implausible, Unrel-Plaus = Unrelated-Plausible, Unrel-Implaus = Unrelated-Implausible.

## 3. Results

### 3.1. Behavioral results

Mean accuracy for the comprehension questions was high in all conditions, indicating that participants were attending to and comprehending the experimental sentences. In the Related-Plausible condition accuracy was 94%, in the Related-Implausible 95%, in the Unrelated-Plausible 93%, and in the Unrelated-Implausible 95%.

### 3.2. ANOVAs on ERPs

Fig. 2 shows the grand-average ERP waveforms time-locked to the target noun for all four conditions on electrode Cz, along with bar graphs showing average voltages in the N400 and P600 time windows for the anterior, central, and posterior electrodes. Table 3 reports the results of the statistical analyses in these time windows.

The analysis in the N400 time window revealed only a main effect of Association, with unrelated targets eliciting a larger negativity (M = −0.98, SD = 2.27) than related ones (M = 1.23, SD = 2.43), which was broadly distributed over the scalp.

In the P600 time window there was a main effect of Plausibility, with implausible targets eliciting a larger positivity (M = 2.96, SD = 2.29) than plausible ones (M = 1.81, SD = 2.87), and an interaction of Association and AP distribution, likely reflecting an influence of component overlap. While the unrelated targets are more negative than the related targets in fronto-central electrodes, continuing the N400 effect ($M_{UNRELATED}$ = 1.72, $M_{RELATED}$ = 2.12), they become more

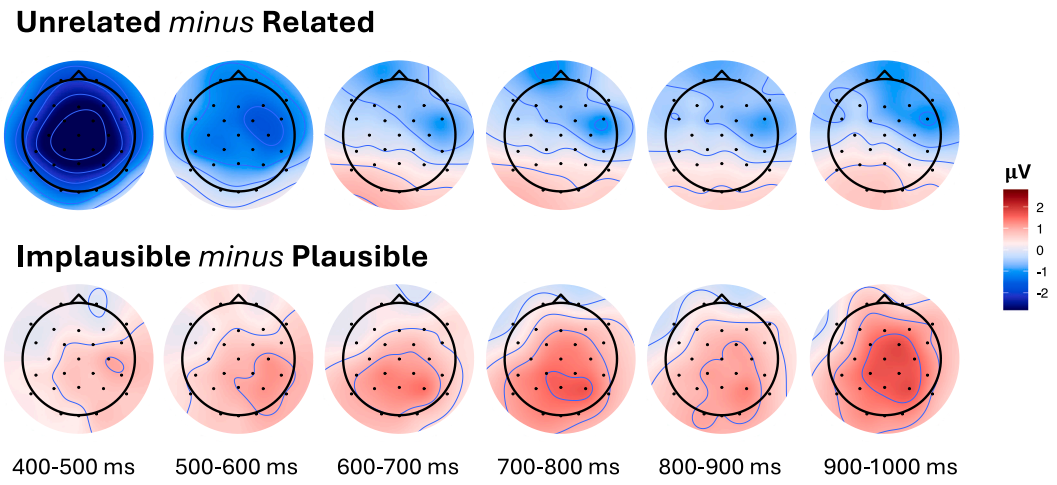



**Fig. 3.** Topographic maps of the effects of Association (top) and Plausibility (bottom) in six consecutive time windows.

**Table 3**
Results of the ANOVAs on ERPs to target nouns across the N400 time window (300–500 ms) and the P600 time window (600–1000 ms).

| | | N400 | | P600 | |
|---|---|---|---|---|---|
| | *df* | *F* | *p* | *F* | *p* |
| Association | (1, 27) | **38.7** | **< .001** | < 1 | .72 |
| Association × AP | (2, 54) | 1.04 | .36 | **4.31** | **.02** |
| Association × Laterality | (2, 54) | 1.15 | .32 | 1.47 | .24 |
| Association × AP × Laterality | (4, 108) | < 1 | .92 | < 1 | .79 |
| Plausibility | (1, 27) | 1.72 | .20 | **5.14** | **.03** |
| Plausibility × AP | (2, 54) | 2.12 | .13 | 1.77 | .18 |
| Plausibility × Laterality | (2, 54) | 2.11 | .13 | 1.18 | .31 |
| Plausibility × AP × Laterality | (4, 108) | 1.32 | .27 | < 1 | .96 |
| Association × Plausibility | (1, 27) | < 1 | .53 | < 1 | .97 |
| Association × Plausibility × AP | (2, 54) | < 1 | .91 | 2.68 | .08 |
| Association × Plausibility × Laterality | (2, 54) | < 1 | .99 | < 1 | .60 |
| Association × Plausibility × AP × Laterality | (4, 108) | < 1 | .87 | < 1 | .44 |

Notes. AP = Anterior–Posterior distribution.

positive in posterior electrodes ($M_{UNRELATED}$ = 3.46, $M_{RELATED}$ = 3.17), where the effect is presumably driven by the large positivity elicited by implausible targets.

The topographic maps of the effects of Association and Plausibility across six consecutive time windows starting at 400 ms are shown in Fig. 3. The N400 effect for unrelated compared to related targets lingers until 600 ms. In later time windows, unrelated targets become more positive at posterior sites (see also the results of the ANOVAs in Appendix B). The effect of Plausibility emerges around 600 ms at posterior sites and becomes more broadly distributed around 700 ms. The time-window analysis also shows a significant interaction of Association, Plausibility, and AP distribution in the 500–600 ms window (along with a main effect of Association) and in the 700–800 ms window (along with a main effect of Plausibility; see Fig. C.7 in Appendix C). These interactions reveal an emerging posterior positivity for unrelated and implausible targets already at 500 ms, with the plausibility effect in later time windows being distributed over posterior electrodes for unrelated targets, while more broadly distributed (and slightly more frontal) for related targets.

Finally, we conducted an ANOVA comparing the Unrelated-Implausible and the Related-Plausible conditions – the key contrast for evaluating component overlap. In DBC19 this contrast revealed an N400 effect only, which was argued to obscure the P600 effect of implausibility (Brouwer, Delogu, & Crocker, 2021; Delogu et al., 2021). As can be seen in Fig. 4, the results of the present experiment revealed a biphasic N400-P600 effect. In the N400 time window, Unrelated-Implausible targets elicited a larger N400 compared to Related-Plausible ones, F(1, 27) = 12.787, p = 0.001 (Fig. 4, left). In the P600 time window, there was an interaction of Plausibility × AP distribution, F(2, 54) = 5.488, p = 0.01, indicating that Unrelated-Implausible targets were more positive than Related-Plausible ones in posterior electrodes (Fig. 4, right).

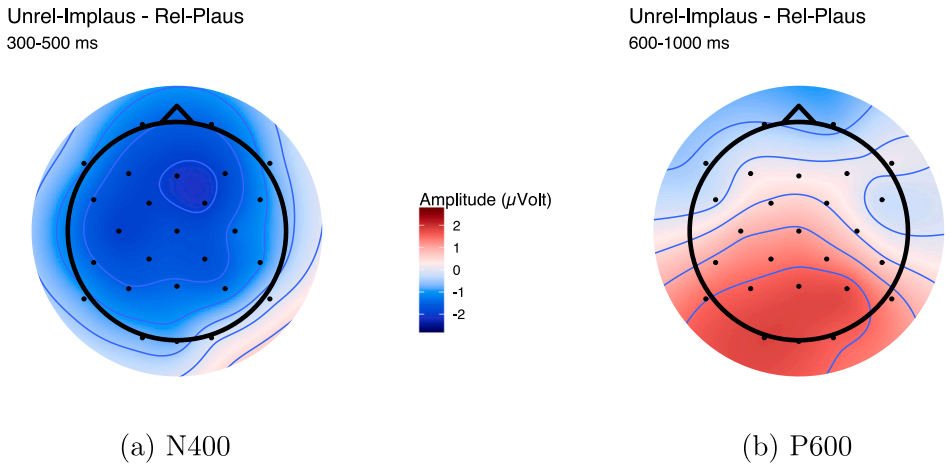
To summarize, the analysis of the ERP data revealed an N400 effect driven by association and a P600 effect driven by plausibility, consistent with previous findings (DBC19, DBC21) and supporting single-stream models such as the RI account (Brouwer et al., 2017, 2012). Additionally, the P600 effect of implausibility appears more pronounced over posterior electrodes for unrelated items, while more broadly distributed for related ones. Importantly, we observed a biphasic N400-P600 effect for Unrelated-Implausible targets compared to Related-Plausible controls, suggesting a mitigation of component overlap in this experiment compared to previous studies in which similar contrasts resulted in an N400 effect only (e.g., DBC19).

### 3.3. rERP analysis

We first assessed a model that includes target word plausibility and association as continuous predictors, using the ratings collected in the norming studies. The model specification for the rERP analysis is as follows:

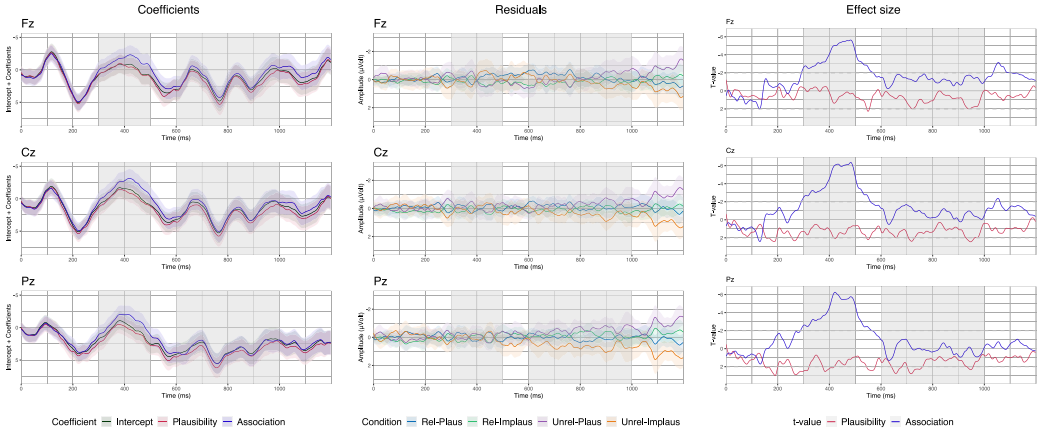
**Model 1:** $Y = \beta_0 + \beta_1 \text{Association} + \beta_2 \text{Plausibility} + \epsilon$

Fig. 5 displays the model's coefficients (anchored to the intercept), the residual error (averaged by condition) and the t-values over time. While the coefficients for association predict more negative-going voltages in the N400 time window, resulting in the N400 effect observed for unrelated versus related targets in ERPs, the coefficients for plausibility predict generally more positive-going voltages for more implausible targets, particularly in centro-posterior electrodes.




(a) N400 (b) P600

**Fig. 4.** Topographic maps of the N400 (left) and P600 (right) effects elicited by Unrelated-Implausible compared to Related-Plausible targets.




(a) Coefficients (b) Residuals (c) t-values

**Fig. 5.** Regression coefficients, residual error, and t-values (based on across-subjects rERPs) at midline electrodes from Model 1.

The residual error by condition, however, indicates that the model does not fully capture the effects in the later time window, with particularly high residuals for unrelated targets in posterior electrodes after 800 ms. This suggests that the plausibility effect observed for unrelated items is larger than what the plausibility ratings predict. As can be seen in Table 2 (here repeated for convenience as Table 4) and Fig. 1, the difference in plausibility between plausible and implausible items is smaller for unrelated targets than for related ones, while the P600 effects show the opposite pattern, indicating that association and plausibility may interact in the late P600 time window. This is consistent with the results of the ANOVAs on the P600 time window (significant in the 700–800 ms time window), suggesting the presence of an interaction between Association and Plausibility driven by a larger P600 effect elicited by Unrelated-Implausible targets in posterior sites as compared to the P600 effect elicited by Related-Implausible targets at the same sites.

**Table 4**
Results of the three norming studies. Mean scores are reported by condition, with standard deviations in parentheses.

| Condition | Related-Plausible | Related-Implausible | Unrelated-Plausible | Unrelated-Implausible |
|---|---|---|---|---|
| Cloze Probability | 0.50 (0.23) | 0.04 (0.06) | 0.004 (0.02) | 0.00 (0.00) |
| Association Rating | 6.51 (0.49) | 6.51 (0.49) | 1.70 (0.63) | 1.70 (0.63) |
| Plausibility Rating | 6.38 (0.62) | 1.97 (0.62) | 5.03 (0.78) | 2.14 (0.75) |

We, therefore, fitted a model including both predictors and their standardized interaction:

$$\textbf{Model 2}: \quad Y = \beta_0 + \beta_1 \text{Association} + \beta_2 \text{Plausibility} + \beta_3 \text{Association} \times \text{Plausibility} + \epsilon$$

As can be seen in the top-left panel of Fig. 6, this model reduces the residual error in the late positive time window relative to Model 1 (Fig.

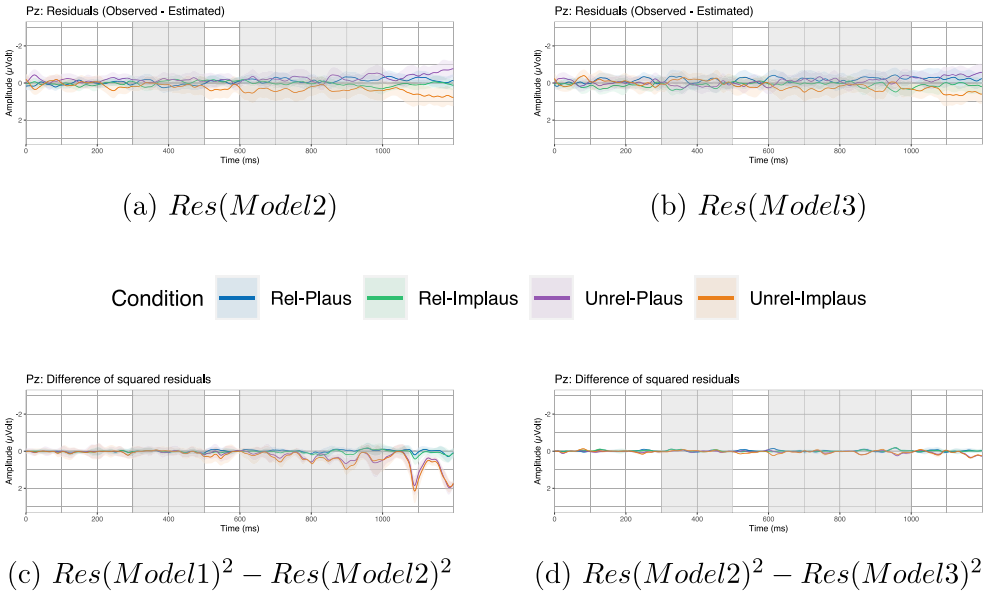



**Fig. 6. Top**: Residual errors from Model 2 (left) and Model 3 (right) at electrode Pz; **Bottom**: Difference of squared residuals between Model 1 and Model 2 (left) and between Model 2 and Model 3 (right).

5). To better visualize this improvement, we computed the difference between squared residuals of Model 1 and Model 2 (Model 1 - Model 2). A positive difference indicates that Model 2 has smaller squared residuals than Model 1. The bottom-left panel of Fig. 6 shows that, particularly in the late part of the segment, the model including the interaction indeed has lower residuals than Model 1. To further assess model fit, we computed the adjusted $R^2$ for both models at the three midline electrodes. The adjusted $R^2$, averaged across time slices and subjects, is higher for Model 2 (Fz: 0.0037; Cz: 0.0030; Pz: 0.0052) than for Model 1 (Fz: 0.0014; Cz: 0.0025; Pz: 0.0044) indicating that Model 2 provides a better fit to the data.

One possible interpretation of this interaction relates to a specific consequence of the design: the Unrelated-Implausible condition shares the same context sentence as the Related-Plausible one (e.g., 'Mark sat down in the restaurant. Before long he opened the…'), likely generating an expectation for the Related-Plausible target (e.g., *menu*). The cloze probability of *menu* is .5 in the Related-Plausible condition (and .04 in the Related-Implausible condition) while the cloze probability of *house door* is extremely low in both Unrelated conditions (see Table 1 and Table 4). When participants encountered *house door* instead of *menu* in the Unrelated-Implausible condition, their predictions were disconfirmed, potentially enhancing the P600 effect. This suggests that the cost of integrating the implausible target word is further increased when an alternative interpretation is highly expected. If this hypothesis is correct, then the stronger the expectation for the related target (*menu*) in the Related-Plausible condition, the stronger the P600 effect elicited by the Unrelated-Implausible target should be.

To investigate this hypothesis, we fitted an rERP model that included Association, Plausibility and the per item cloze probability of the related target word depending on the context [Related-Plausible: Mean = .50 (SD = .23); Related-Implausible: .04 (.06); Unrelated-Plausible: .04 (.06); Unrelated-Implausible: .50 (.23), see Tables 1 and 4]:

$$\textbf{Model 3}: \quad Y = \beta_0 + \beta_1 \text{Association} + \beta_2 \text{Plausibility} + \beta_3 \text{RelatedTargetCloze} + \epsilon$$

As can be seen in the top-right panel of Fig. 6, the residual error for this model improves compared to the simple model, similar to the model including the interaction term. Accordingly, the difference of squared residuals between Model 2 (with the interaction) and Model 3 is close to 0 (Fig. 6, bottom-right panel). This finding suggests that the larger late P600 effect to Unrelated-Implausible targets may reflect the additional cost of integrating an implausible word in contexts where there is a highly expected alternative interpretation (see also, e.g., Aurnhammer et al., 2023; DeLong et al., 2014; Kuperberg et al., 2020; Van Petten & Luka, 2012).

To summarize, the rERP analysis revealed that the amplitude of the N400 is sensitive to association but not to plausibility, while the P600 is sensitive to plausibility. Consistent with the ERP analysis, we found an interaction of association and plausibility in the P600 time window, resulting in a larger P600 effect in centro-posterior electrodes for unrelated targets. The regression analysis shows that this effect cannot be solely attributed to the degree of implausibility of the targets, but it likely also reflects an additional cost associated with disconfirmed predictions, since *menu* is more expected than *house door* following a context sentence like 'John entered the restaurant. Before long he opened the… '. However, this aspect of the experimental design was also present in DBC21, where no biphasic effect for unrelated implausible targets compared to related controls was observed. Thus, it is unlikely that this factor plays a crucial role in enhancing the P600 compared to DBC21. The key factor appears to be the task itself, specifically the use of comprehension questions rather than plausibility judgments.

## 4. Discussion

A well-documented finding in the ERP literature is that the processing of incongruent linguistic input typically modulates the N400 component. However, the extent to which these modulations are accompanied by a P600 response varies considerably, with the factors underlying this variability remaining unclear (see Van Petten & Luka, 2012). Brouwer and Crocker (2017) hypothesized that spatiotemporal overlap between ERP components (Luck, 2005)—an often-overlooked

inherent property of the ERP signal—may account for these inconsistencies. They argue that the extent to which the P600 effect survives overlap with the N400 in the observed signal depends on the relative magnitude of the latent components, which may in turn depend on how deeply the linguistic input is processed for comprehension. Based on evidence that the P600 is sensitive to task demands (e.g., Kolk et al., 2003; Leckey & Federmeier, 2020; Schacht et al., 2014), Brouwer and Crocker predict that deeper comprehension enhances the P600, increasing its likelihood of surviving component overlap. The aim of the current study was to investigate the N400-P600 dynamics when readers are asked comprehension questions rather than plausibility judgments as used in prior studies showing component overlap (DBC19, DBC21). Like in those studies, we manipulated whether a word was related or unrelated to its preceding context and whether it was plausible or implausible.

Two key findings emerge from our results: first, we replicated DBC19 and DBC21 with respect to the N400's sensitivity to lexical association (and not plausibility) and the P600's sensitivity to plausibility, further establishing that these effects are robust across different tasks (see also Aurnhammer et al., 2023, 2021). Additionally, however, despite the presence of an N400 effect, we observed a P600 effect for unrelated implausible targets compared to the related plausible controls, which was not observed in DBC19. This observed biphasic N400-P600 pattern suggests a mitigation of component overlap when an engaging task is employed. Both of these findings have important implications for the theoretical understanding of the N400 and P600, strongly supporting single-stream accounts such as the Retrieval-Integration (RI) model (Brouwer et al., 2017, 2012) over competing theories, as we will argue below.

In the present study, the P600 is elicited in response to implausibilities that cannot be resolved through syntactic or semantic reanalysis, such as thematic role reassignment in so-called reversal anomalies (e.g., van Herten et al., 2006, 2005; Hoeks et al., 2004; Kim & Osterhout, 2005; Kolk et al., 2003; Kuperberg et al., 2003). We observed a P600 effect for implausible targets in well-formed sentences without thematic role violations, replicating previous results (DBC19, DBC21). This challenges theories of the P600 as indexing syntactic/semantic reanalysis or conflict detection between distinct processing streams (e.g. Bornkessel-Schlesewsky & Schlesewsky, 2008; van Herten et al., 2006, 2005; Kim & Osterhout, 2005). More generally, our results are difficult to reconcile with multi-stream models that predict either a monophasic N400 or P600 effect in response to incongruent input, depending on whether a more plausible alternative interpretation is available (P600) or not (N400) (e.g., Kim & Osterhout, 2005; Ryskin et al., 2021). Overall, the existing evidence indicates that the P600 reflects a default brain response to processing (implausible) linguistic input and that the variability in observing a P600 effect as part of a biphasic pattern may result from spatiotemporal overlap between the N400 and P600.

Second, our findings indicate that the task may influence the magnitude, but not the occurrence of the P600. The P600 effect for implausibility is consistently observed with comprehension questions (in the present study), plausibility judgments (DBC19, DBC21) and even in the absence of an overt task (Nieuwland & van Berkum, 2005; Van Petten & Luka, 2012). Importantly, different tasks may differentially shape the way comprehenders process a text. Plausibility judgments might encourage readers to adopt a more strategic approach while reading, focusing narrowly on plausibility without fully engaging with the content of the text. In contrast, unpredictable comprehension questions – relative to both their occurrence and focus – require readers to pay more attention to the entire text, promoting more natural reading and deeper processing.

In summary, we found that the N400 is larger for unassociated targets and the P600 for implausible ones, and that the extent to which the P600 for implausible *and* unassociated targets emerges in a biphasic N400-P600 pattern in the scalp-recorded signal may depend on the reader's goals and depth of comprehension. Consistent with our findings, previous studies have reported biphasic N400-P600 effects in response to contextually unassociated semantic anomalies when comprehension questions on a proportion of trials are used (e.g., DeLong & Kutas, 2020; Klingvall & Heinat, 2024). Indeed, biphasic effects have also been observed with a plausibility judgment task for strong violations such as animacy and selectional restriction violations (Aurnhammer et al., 2021; Brothers et al., 2020; Kuperberg et al., 2020), as opposed to the more subtle script-based plausibility manipulation investigated here and in DBC19. Due to spatiotemporal component overlap with the N400 (Brouwer & Crocker, 2017; Delogu et al., 2021), however, the presence of a P600 effect in the scalp-recorded signal is contingent upon the strength of the plausibility manipulation (e.g., semantic violations vs. script-based plausibility), the richness of the context (Brothers et al., 2020; Kuperberg et al., 2020), and the depth of comprehension (Brothers et al., 2020; Schacht et al., 2014; Wang et al., 2024).

Overall, these findings strongly support single-stream accounts such as the RI theory (Brouwer et al., 2017, 2012; Venhuizen & Brouwer, 2025). Specifically, this account posits that language processing proceeds through (partially overlapping) cycles of retrieval (N400) and integration (P600) processes: the N400 is modulated by factors such as lexical association and contextual expectancy that determine the ease of accessing word meaning from long-term memory (Aurnhammer et al., 2021; Kutas & Federmeier, 2000, 2011), while the P600 indexes integration effort due to semantic, syntactic or pragmatic factors (Brouwer, Delogu, Venhuizen, & Crocker, 2021; Brouwer et al., 2012). When both retrieval and integration effort increase, the theory predicts a biphasic N400-P600 response. However, when retrieval and integration processes (partially) overlap, their corresponding indexing components do as well, which can attenuate post-N400 positivities, resulting in the absence of a P600 effect. Crucially, the theory predicts integration effort, reflected in the P600, to be modulated by depth of comprehension: deeper processing results in enhanced P600 amplitude, increasing the chances of the P600 to survive spatiotemporal overlap with the N400.

Our findings have also important methodological implications, particularly concerning the role of the experimental task. The presence or absence of a 'post-N400 positivity' may depend on how readers approach a text, which is strongly shaped by task demands. Without careful consideration of the properties of the task, the reading strategies they may induce, and the cognitive processes they may affect – along with the ERP components they are indexed by – the interpretation of observed ERP effects may be misguided, leading to flawed theoretical conclusions about their functional significance.

## 5. Conclusions

We conducted an ERP study to investigate the extent to which previously observed spatiotemporal overlap between the N400 and P600 components is mitigated when readers are required to perform a more ecological task encouraging deeper comprehension. We found that, when using comprehension questions, the P600 effect elicited by implausible words survives overlap with the N400, resulting in a biphasic pattern. Our findings provide strong evidence for the biphasic nature of the N400-P600 complex, supporting single-stream accounts of language comprehension, emphasizing the importance of carefully considering task demands and spatiotemporal component overlap when interpreting the presence or absence of ERP effects in the observed signal.

## CRediT authorship contribution statement

**Francesca Delogu:** Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Christoph Aurnhammer:** Writing – review & editing, Software, Methodology,

Formal analysis. **Harm Brouwer:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Matthew W. Crocker:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 232722074 – SFB 1102. The authors would like to thank Miriam Schulz for help creating the stimuli.

## Appendix A. Experimental materials

Each item consists of a context sentence and a target sentence in four conditions: *Related-Plausible*, *Related-Implausible*, *Unrelated-Plausible*, and *Unelated-Implausible* (see Table 1 in the Introduction). The context sentence differs between the Plausible and Implausible conditions typically in the verb. The target sentence differs between the Related and Unrelated conditions in the target word. At least three words following the target word remain the same across all conditions. The four conditions are represented as follows:

Mark [**erreichte/verließ**] den Spielplatz. Fröhlich lief er zu der [Schaukel/Eisdiele] und von dort aus winkte er seinen Eltern.

The four conditions can be distinguished as follows:

- **Related-Plausible**: Mark **erreichte** den Spielplatz. Fröhlich lief er zu der Schaukel und von dort aus winkte er seinen Eltern.
- **Related-Implausible**: Mark **verließ** den Spielplatz. Fröhlich lief er zu der Schaukel und von dort aus winkte er seinen Eltern.
- **Unrelated-Plausible**: Mark **verließ** den Spielplatz. Fröhlich lief er zu der Eisdiele und von dort aus winkte er seinen Eltern.
- **Unrelated-Implausible**: Mark **erreichte** den Spielplatz. Fröhlich lief er zu der Eisdiele und von dort aus winkte er seinen Eltern.

The full set of materials is reported below in the reduced form, along with the comprehension questions.

1. Mark [**erreichte/verließ**] den Spielplatz. Fröhlich lief er zu der [Schaukel/Eisdiele] und von dort aus winkte er seinen Eltern.
   Q: Hat Mark seine Eltern gesehen?
2. Leo [**ging/kam**] zum Strand. Eine Weile spielte er in dem [Sand/Park] und danach lief er zufrieden nach Hause.
3. Clara [**begann/beendete**] die Opernaufführung. Endlich betrat sie die [Bühne/Bahn] und plötzlich war sie sehr nervös.
4. Katrin [**betrat/verließ**] das Nagelstudio. Spontan entschied sie sich für eine [Maniküre/Spritztour] mit dem neuen [Nagellack/Wagen].
   Q: Ist Katrin beim Friseur gewesen?
5. Tim war [**dabei/fertig damit**], den Müll zu entsorgen. Prompt öffnete er den [Eimer/Comic] und schon entdeckte er [die Maden/er den Plot-Twist].
6. Christina [**begann mit dem/kam vom**] Babysitten. Nach einer Weile wechselte sie die [Windel/Buslinie] und kurz darauf bereitete sie [den Babybrei zu/ihre Haltestelle.].
7. Paul [**fing/war**] mit der Reparatur [**an/fertig**]. Sofort nahm er den [Schraubenschlüssel/Unterhaltungsroman] aus dem verstaubten [Werkzeugkasten/Bücherregal].
   Q: Nahm Paul etwas aus dem Kleiderschrank?
8. Thomas [**begann damit/war damit fertig**], die Pflanzen zu versorgen. Fröhlich suchte er die [Gießkanne/Teetasse] und auf der Stelle [goss er die Yuccapalme/machte er einen Kräutertee.].
9. Johann [**fing mit den/hatte die**] Rechenaufgaben [**an/erledigt**]. Gelassen holte er seinen [Taschenrechner/Videorecorder] aus dem Fach im [Schulranzen/Regal].
10. Olaf ging zum Bräunen [**hinaus/herein**]. Sofort öffnete er die [Sonnencreme/Arbeitsmappe] und dann begann er [sich einzucremen/zu arbeiten].
    Q: War Olaf zum Bräunen draußen?
11. Susi [**begann/war fertig**] mit der Katzenfütterung. Schnell öffn ete sie die [Dose/Cola] und dann holte sie [den Futternapf aus dem Schrank/die Eiswürfel aus dem Gefrierfach].
12. Anna [**betrat/verließ**] das Fitnessstudio. Sofort ging sie zu der [Hantel/Vespa] und daraufhin machte sie [ein paar Wiederholungen/sich auf den Weg].
13. Elena [**ging hinaus in den** Sturm/**kam aus dem** Sturm **herein**]. Schnell öffnete sie ihren [Schirm/Kamin] und dabei verwünschte sie das schlechte Wetter.
    Q: Freute Elena sich über das Wetter?
14. Adam [**wollte ein /hatte das**]Experiment durchführen. Rasch ging er los zum [Labor/Golf] und dort traf er seine [Kollegin Frau].
15. Mia [**ging zum/kam vom**] Turnen. Als Erstes machte sie eine [Dehnübung/Kürbissuppe] und danach [die erste Turnübung/den leckeren Nachtisch].
16. Harald [**war beim/kam vom**] Ski. Fröhlich setzte er sich auf den [Lift/Hocker] und dabei wurde er ganz müde.
    Q: Ist Harald Snowboard gefahren?
17. Paulina [**betrat/verließ**] den Pub. Direkt ging sie zu dem [Barkeeper/Chauffeur] und lächelnd begrüßte sie ihn.
18. Diana [**wollte/hatte**] Wasser [**trinken/getrunken**]. Rasch suc hte sie ein [Glas/Handy] und dann öffnete sie den [Hahn/Chat].
19. Jonas [**fing damit an/war damit fertig**], die Bratkartoffeln zu machen. Vorsichtig öffnete er das [Öl/Bier] und schnell goss er es in [die Bratpfanne/den Bierkrug].
    Q: War Jonas beim Öffnen vorsichtig?
20. Marco hat mit dem Putzen [**begonnen/aufgehört**]. Direkt schn appte er sich den [Lappen/Roller] und dann begann er [zu schrubben/loszufahren].
21. Thea [**begann/war fertig**] damit, das Gras zu stutzen. Auf der Stelle startete sie den [Rasenmäher/Radioapparat] und dabei pfiff sie eine Melodie.
22. Petra [**ging zu/kam von**] der Beerdigung. Bedrückt betrat sie den [Friedhof/Marktplatz] und dort begrüßte sie [die Trauergemeinde/den Gemüseverkäufer].
    Q: Hat Petra jemanden begrüßt?
23. Marian [**wollte/hatte**] ein Kleid [**kaufen/gekauft**]. Beschwingt betrat sie die [Boutique/Taverne] und sofort probierte sie [ein hübsches Kleid an/einen neuen Drink aus].
24. Nora [**bereitete sich auf das** Motorradfahren **vor/kam vom** Motorradfahren **zurück**]. Zuerst suchte sie ihren [Helm/Krimi] und schließlich fand sie ihn im Flur.
25. Samantha [**ging los/kam vom**] zum Poker. Unverzüglich betrat sie das [Casino/Kaufhaus] und zuerst ging sie zu den [Automaten/Spielsachen].
    Q: Hat Samantha Poker gespielt?
26. Ruth [**war beim/kam vom**] Boxtraining. Wütend trat sie gegen den [Sack/Stamm] und dabei tat sie sich weh.

27. Julian [**bereitete die/kam von der**] Demonstration vor. Emsig beschrieb er das [Schild/Tagebuch] mit den besonderen [Parolen der Partei/Ereignissen des Tages].
28. Maja [**war beim/hörte mit**] dem Fechten [/**auf**]. Unverzüglich nahm sie den [Degen/Snack] aus der Tasche heraus. Q: Nahm Maja etwas aus der Tasche?
29. Tom hatte sein Frühstück [**begonnen/beendet**]. Sogleich schnitt er das [Brot/Holz] und direkt danach [nahm er die Marmelade/befeuerte er den Holzofen].
30. Lars [**eilte zum/schwänzte den**] Gottesdienst. Letztendlich ging er in die [Kirche/Natur] und dort setzte er sich auf eine Bank.
31. Lisa [**rannte zum/schwänzte den**] Schulunterricht. Schließlich betrat sie das [Klassenzimmer/Stadtzentrum] und als Erstes [entschuldigte sie sich für die Verspätung/ging sie in das Shoppingzentrum].
    Q: War Lisa bei der Nachhilfe?
32. Mark [**betrat/verließ**] das Kino. Munter kaufte er ein [Ticket/Hemd] für nur sieben Euro.
33. Manuel [**betrat/verließ**] die Bankfiliale. Entschlossen ging er zu dem [Schalter/Zirkus] und dort erkundigte er sich nach den [möglichen Krediten/aktuellen Preisen].
34. Emilia [**kam in die/ging aus der**] Disko [**hinein/hinaus**]. Nach kurzem Überlegen ging sie zu der [Bar/Villa] und nachdenklich betrachtete sie [die Cocktailauswahl/den Vorgarten].
    Q: Ist Emilia in der Disko gewesen?
35. Felix [**begann/hörte**] dem Zeichnen [/**auf**]. Munter nahm er den [Stift/Bass] in die Hand und [skizzierte ein Porträt/übte einen Song].
36. Sarah [**betrat/verließ**] das Freibad. Fröhlich ging sie zu der [Rutsche/Gaststube] und schon bald [sauste sie hinunter/bestellte sie etwas].
37. Heinrich [**wollte/hatte**] einen Rosenstrauß[**kaufen/gekauft**]. Nach kurzer Überlegung ging er zu dem [Blumenladen/Feinkostladen] und dort kaufte er zehn [Rosen/Pralinen].
    Q: Hat Heinrich etwas im Bäckerladen gekauft?
38. Lena [**bereitete sich darauf vor/hatte damit aufgehört**], ein Brettspiel zu spielen. In Ruhe las sie die [Anleitung/Biographie] und nach einer halben Stunde war sie fertig.
39. Elisa [**ging zu /kam von**] der Wahrsagerin. Skeptisch blickte sie auf die [Kugel/Wiese] und dabei dachte sie über ihre Zukunft nach.
40. Egon [**fing damit an/hörte damit auf**], den Himmel zu beoba chten. Nach einem Moment entdeckte er einen [Stern/Ball] und schnell rief er seinen Bruder.
    Q: Hat Egon einen Bruder?
41. Paula [**ging zum/kam vom**] Badesee. Direkt schlüpfte sie in ihren [Bikini/Pyjama] und als Nächstes [sprang sie in den See/pu tzte sie ihre Zähne].
42. Bruno [**begann/war fertig**] mit dem Händewaschen. Schnell nahm er die [Seife/Birne] aus der Schale und [wusch seine Hände/schnitt sie in Stücke].
43. Dominik [**ging hinaus in die/kam herein aus der**] Kälte. Prompt rutschte er aus auf dem [Glatteis/Laminat] und sofort fiel er hin.
    Q: War es draußen kalt?
44. Leonie [**ging zum/kam vom**] Mittagessen. Munter spazierte sie zu der [Mensa/Weide] und dort setzte sie sich an [einen Tisch/eine Bank].
45. Isabel [**spazierte durch das/kam aus**] dem Museum. Kurze Zeit betrachtete sie das [Gemälde/Gewässer] und dann ging sie zu [den Statuen/ihrem Auto].
46. Sabine [**hatte sich eine Grippe eingefangen/von der Grippe erholt**]. Am Vormittag betrat sie die [Arztpraxis/Turnhalle] und dort begrüßte sie die [Ärztin/Trainerin].
    Q: War Sabine krank?
47. Rolf [**musste/hatte**] einen neuen Wecker [**kaufen/gekauft**]. Daraufhin ging er zu einem [Uhrmacher/Fischmarkt] und spontan kaufte er einen [Wecker mit Tageslichtfunktion/Rotbarsch für das Abendessen].
48. Kim hatte das Hotel [**betreten/verlassen**]. Zuerst ging sie zu der [Rezeption/Apotheke] und kurz sprach sie mit dem [Empfangsmitarbeiter/Apotheker].
49. Manfred [**erreichte/verließ**] das Theater. Wenig später war er in dem [Saal/Pool] und kurz darauf kam auch seine Freundin dazu.
    Q: Hat Manfred sich mit seiner Freundin getroffen?
50. Johannes war [**zum/vom**] Bergsteigen [**gegangen/zurückge kommen**]. Am Nachmittag erreichte er den [Gipfel/Kongress]
51. Augusta [**ging in die/kam von der**] Reitschule. Endlich setzte sie sich in den [Sattel/Sessel] und sofort begann sie [loszureiten/einzunicken].
52. Gerhard hatte den Feuerwehreinsatz [**begonnen/beendet**]. Prompt nahm er den [Schlauch/Weißwein] in die Hand und [richtete ihn auf den Brandherd/öffnete ihn mit dem Korkenzieher].
    Q: Ist Gerhard Feuerwehrmann?
53. Jenni [**wollte/kam**] mit dem Hund [**rausgehen/nach Hause**]. Eilig suchte sie die [Leine/Pizza] und dann wartete sie ungeduldig, bis [der Hund kam/der Ofen heißwar].
54. Franziska [**war in die/kam von der**] Einkaufsstraße [**gegan gen/zurück**]. Gedankenverloren betrachtete sie das [Schaufenster/Kinderzimmer] und nach einem Moment [ging sie in das Geschäft]/räumte sie das Spielzeug auf.
55. Lennox [**erreichte den/kam vom**] Supermarkt [/**zurück]**. Als Erstes nahm er sich einen [Einkaufswagen/Staubsauger] und als Nächstes [betrat er den Supermarkt/saugte er den Flur].
    Q: War Lennox einkaufen?
56. Daniel [**ging in den/kam vom**] Wald [/**heim**]. Auf der Stelle stolperte er über eine [Wurzel/Gitarre] auf dem Boden und ärgerte sich.
57. Frauke [**wollte eine/hatte ihre**] Yogasession [**machen/beendet**]. Munter öffnete sie ihre [Matte/Banane] und sofort begann sie, zu [meditieren/essen].
58. Karsten [**betrat/verließ**] das Wohnzimmer. Sogleich ging er zu dem [Sofa/Date] und dort wurde ihm langweilig.
    Q: Hat Karsten sich amüsiert?
59. Edith [**wollte einen/hatte ihren**] neuen Personalausweis [**beantragen/abgeholt**]. Daraufhin machte sie den [Termin/Haushalt] für den nächsten Tag.
60. Leon [**begann/beendete**] seinen Mittagsschlaf. Sogleich hatte er einen [Traum/Streit] über ein besonderes [Raumschiff mit Aliens/Spielzeug mit seiner Schwester].
61. Maren [**brauchte/hatte**] einen neuen Haarschnitt [/**bekommen**]. Nach kurzer Überlegung betrat sie den [Friseursalon/Bäckerladen] und dort ließsie sich [einen modischen Schnitt verpassen/ein leckeres Teilchen einpacken].
    Q: Hat Maren den Supermarkt betreten?
62. Sina [**ging in/verließ**] das Callcenter. Als Nächstes machte sie einen [Anruf/Ausflug] und dabei schaffte sie es, [den Kunden zu überzeugen/zehn Kilometer zu gehen].
63. Markus [**setzte sich im/verließdas**] Restaurant. Bald darauf öffnete er die [Speisekarte/Wohnungstür] und als Nächstes [wählte er sein Gericht/begüßte er seine Kinder].
64. Nils [**wollte/hatte**] die Fahrbahn [**überqueren/überquert**]. Schnell ging er zu der [Ampel/Disco] und dort wartete er, bis [sie grün wurde/er hineingelassen wurde].
    Q: Ist Nils langsam gegangen?

65. Sandra [**wollte/war**] mit dem Frisieren [**anfangen/fertig**]. Sogleich holte sie den [Kamm/Sekt] aus dem Regal und [frisierte sich/ging zur Party].
66. Philipp [**musste/hörte**] mit dem Fegen [**anfangen/auf**]. Entschlossen nahm er den [Besen/Rum] in die Hand und [fegte die Wohnung/trank einen Schluck].
67. Maria [**ging in/verließ**] das Badezimmer. Als Nächstes putzte sie die [Dusche/Terrasse] mit dem neuen Putzmittel.
    Q: Hat Maria das neue Putzmittel benutzt?
68. Robert [**musste/hatte**] eine Durchsage [**machen/beendet**]. Sofort suchte er das [Mikrofon/Gasthaus] und dann las er [den Text vor/er die Empfehlungen des Tages].
69. Maike [**ging in/verließ**] das Wellnesscenter. Unverzüglich bekam sie eine [Massage/Erkältung] und danach hatte sie keine [Verspannungen mehr/Energie mehr].
70. Theodor schritt [**in den/aus dem**] Operationssaal [**hinein/**]. Kurz darauf griff er nach dem [Skalpell/Butterbrot] und hektisch begann er den [Eingriff/zu essen].
    Q: War Theodor im Operationssaal?
71. Lea [**wollte ein/ war fertig**] mit ihrem Graffiti machen. Behutsam öffnete sie die [Spraydose/
    Chipstüte] und nach ein paar Minuten war [das Graffiti schon fertig/war die Tüte schon leer].
72. Ayla [**öffnete/schloss**] das Horoskop. Eilig suchte sie ihr [Sternzeichen/Taschentuch] und wie erwartet [war es negativ/musste sie weinen].
73. Samuel [**wollte in/hatte**] sein Bett [**gehen/verlassen**]. Nach einer Weile schlüpfte er in seinen [Schlafanzug/Sonntagsanzug] und dann machte er [das Licht aus/Spiegeleier].
    Q: Hat Samuel sich angezogen?
74. Ludwig [**betrat/verließ**] die Kapelle. Langsam näherte er sich dem [Altar/Kanal] und dort sah er [den Priester/einen Jogger].
75. Charlotte [**betrat/verließ**] den Schmuckladen. Dann kaufte sie eine [Kette/Frucht] von ihrem letzten Geld.
76. Tobias [**wollte/hörte auf,**] seine Lieblingssendung schauen. Eilig suchte er die [Fernbedienung/
    Sonnenbrille] und fand sie schließlich im Zimmer seines Bruders.
    Q: Suchte Tobias etwas?
77. Alina [**begann/war fertig**] damit, ihren Pony zu kürzen. Rasch griff sie nach der [Schere/Jacke] und dann ging sie [zum Spiegel/hinaus zu dem Treffen].
78. Rosalie [**ging zu/kam von**] ihrer Chorprobe. Lächelnd nahm sie das [Notenblatt/Sonnenbad] und dabei grüßte sie [den Chorleiter/die Nachbarn].
79. Friedrich [**betrat/verließ**] das Hochhaus. Schnellen Schrittes ging er zu dem [Aufzug/Kiosk] und dann fuhr er [in seinen Stock/zum Sport].
    Q: Ist Friedrich in einer Villa gewesen?
80. Anastasia [**begann/beendete**] die Renovierung. Munter suchte sie die [Farbe/Post] für die neue [Wand aus/Mieterin].
81. Kalle [**begann/war fertig**] damit, seinen Bart zu trimmen. Hektisch griff er nach dem [Rasierer/Anorak] und als Nächstes [setzte er ihn an/ging er nach draußen].
82. Eva [**war beim/kam vom**] Schlittschuhlaufen. Fröhlich betrat sie die [Eisfläche/Veranda] und als Erstes [machte sie gekonnt eine Pirouette/setzte sie sich auf die Verandaschaukel].
83. Heiner [**ging los zu/kam von**] der Versteigerung. Munter betrat er das [Auktionshaus/Fotogeschäft] und ziemlich spontan [ersteigerte er dort eine Kamera/kaufte er dort eine Kamera].
84. Aurelia [**erhielt eine/entging einer**] Haftstrafe. Am nächsten Tag betrat sie das [Gefängnis/Rathaus] und letztendlich bekam sie [Angst/eine Entschädigung].
85. Ida [**sollte ein/hatte**] das Interview [**durchführen/
    beendet**]. Hastig holte sie ihren [Notizblock/Müsliriegel] aus dem Rucksack hervor.
    Q: Sollte Ida interviewt werden?
86. Gabriel [**wollte in den/kam aus dem**] Urlaub [**fahren/zurück**]. Am nächsten Morgen ging er zu dem [Reisebüro/Sozialamt] und da ließer sich informieren.
87. Marina [**musste/hatte**] ihr Benzin [**auffüllen/aufgefüllt**]. Sogleich fuhr sie zu der [Tankstelle/
    Buchmesse] und prompt begann sie mit dem [Tanken/Vortrag].
88. Olivia hatte [**Hunger bekommen/ihren Hunger gestillt**]. Auf der Stelle ging sie zu dem [Kühlschrank/Schreibtisch] und entschlossen nahm sie [den Joghurt heraus/Ordner].
    Q: Hatte Olivia Durst?
89. Rebecca [**wollte/war**] mit dem Kochen [**anfangen/fertig**]. Bald darauf ging sie in die [Küche/Uni] und als Erstes öffnete sie den [Küchenschrank/Seminarraum].
90. Noah [**musste einen/verließ**] den Zug [**nehmen/**]. Zügig eilte er zu dem [Bahnsteig/Bezirksamt] und dort wartete er ungeduldig auf den Zug/die Besprechung.
91. Noemi [**fing/war**] mit ihren Geographiehausaufgaben [**an/fertig**]. Behutsam öffnete sie das [Lehrbuch/Getränk] und dann suchte sie [den Atlas/ihre Barbie].
    Q: War Noemi unvorsichtig?
92. Noel [**hatte seine Liegestützen gemacht/sich vor den Lieges tützen gedrückt**]. Am nächsten Tag bekam er [Muskelkater/Missbilligung] von den vielen [Liegestützen/Kurst eilnehmern].
93. Bella [**war beim/kam vom**] Zelten [**/zurück**]. Am Abend setzte sie sich an das [Lagerfeuer/Schlagzeug] und sogleich fing sie an, [Marshmallows zu grillen/einen Rhythmus zu spielen].
94. Frieda [**wurde zum/aus dem**] Krankenhaus [**gefahren/entlassen**]. Eilig betrat sie die [Notaufnahme/
    Telefonzelle] und sofort rief sie [nach einem Arzt/ihre Mutter an].
    Q: War Frieda im Krankenhaus?
95. Kaja [**sollte/hatte**] eine Torte [**kaufen/gekauft**]. Hastig ging sie zu der [Konditorei/Ratssitzung] und fröhlich begrüßte sie [den Konditor und wählte eine Torte/die Ratsmitglieder und verteilte die Torte].
96. Liana hatte den Banküberfall [**begonnen/beendet**]. Flink öffnete sie den [Tresor/Truck] und sogleich begann sie, die Beute herauszuholen.
97. Tyler [**hatte sich auf die/war von der**] Kreuzfahrt [**vorbereitet/zurückgekommen**]. Am nächsten Tag betrat er das [Schiff/Dorf] und schon traf er [die ersten Passagiere auf dem Deck/seine besten Freunde auf dem Dorfplatz].
    Q: Hat Tyler am gleichen Tag jemanden getroffen?
98. Finn [**wollte/hatte**] die Schießerei [**beginnen/beendet**]. Als Nächstes griff er nach seiner [Pistole/
    Zigarette] und dann begann er [loszuschießen/zu rauchen].
99. Josephine [**begann/war fertig**] mit dem Haarewaschen. Als Erstes griff sie nach dem [Shampoo/
    Hörgerät] und danach nach dem Conditioner/Handtuch.
100. Lenny [**stieg auf/verließ**] das Boot. Zielstrebig ging er zu dem [Steuer/Kloster] und dann setzte er [die Segel/ sich hinein].
    Q: War Lenny auf einer Yacht?

## Appendix B. ANOVAs on six consecutive time windows between 400 ms and 1000 ms

See Table B.5.

## Appendix C. Topographic maps

See Fig. C.7.

**Table B.5**
Results of the ANOVAs on ERPs to target nouns across six 100-ms time windows, starting at 400 ms post-stimulus onset.

| | | [400–500] ms | | [500–600] ms | | [600–700] ms | |
|---|---|---|---|---|---|---|---|
| | *df* | *F* | *p* | *F* | *p* | *F* | *p* |
| Assoc | (1, 27) | **58.3** | < **.001** | **9.02** | **.006** | < 1 | .69 |
| Assoc × AP | (2, 54) | 1.26 | .29 | < 1 | .41 | **4.05** | **.03** |
| Assoc × Lat | (2, 54) | 1.44 | .13 | 1.60 | .21 | 1.35 | .27 |
| Assoc × AP × Lat | (4, 108) | < 1 | .84 | < 1 | .74 | < 1 | .73 |
| Plaus | (1, 27) | 1.84 | .19 | 1.98 | .17 | 2.93 | .10 |
| Plaus × AP | (2, 54) | 1.53 | .23 | 2.30 | .12 | **4.37** | **.03** |
| Plaus × Lat | (2, 54) | 2.12 | .13 | 2.31 | .11 | < 1 | .54 |
| Plaus × AP × Lat | (4, 108) | 1.33 | .27 | < 1 | .56 | < 1 | .09 |
| Assoc × Plaus | (1, 27) | 1.25 | .27 | < 1 | .52 | 1.26 | .27 |
| Assoc × Plaus × AP | (2, 54) | 2.25 | .14 | **4.87** | **.03** | 2.95 | .08 |
| Assoc × Plaus × Lat | (2, 54) | < 1 | .85 | < 1 | .62 | 2.31 | .12 |
| Assoc × Plaus × AP × Lat | (4, 108) | < 1 | .67 | < 1 | .75 | < 1 | .85 |
| | | [700–800] ms | | [800–900] ms | | [900–1000] ms | |
| | *df* | *F* | *p* | *F* | *p* | *F* | *p* |
| Assoc | (1, 27) | < 1 | .66 | < 1 | .83 | < 1 | .78 |
| Assoc × AP | (2, 54) | 2.71 | .09 | 1.81 | .18 | **5.59** | **.01** |
| Assoc × Lat | (2, 54) | < 1 | .40 | 1.03 | .36 | 2.00 | .15 |
| Assoc × AP × Lat | (4, 108) | < 1 | .60 | < 1 | .80 | < 1 | .73 |
| Plaus | (1, 27) | **6.90** | **.01** | 2.91 | .10 | **6.61** | **.02** |
| Plaus × AP | (2, 54) | 2.07 | .15 | < 1 | .45 | < 1 | .67 |
| Plaus × Lat | (2, 54) | 1.11 | .33 | < 1 | .42 | 1.60 | .21 |
| Plaus × AP × Lat | (4, 108) | < 1 | .87 | < 1 | .96 | < 1 | .88 |
| Assoc × Plaus | (1, 27) | < 1 | .57 | < 1 | .44 | < 1 | .54 |
| Assoc × Plaus × AP | (2, 54) | **4.25** | < **.05** | 3.41 | .06 | < 1 | .52 |
| Assoc × Plaus × Lat | (2, 54) | < 1 | .78 | < 1 | .75 | < 1 | .78 |
| Assoc × Plaus × AP × Lat | (4, 108) | 1.15 | .33 | 1.39 | .25 | 1.37 | .26 |

Notes. Assoc = Association; Plaus = Plausibility; AP = Anterior–Posterior distribution; Lat = Laterality.

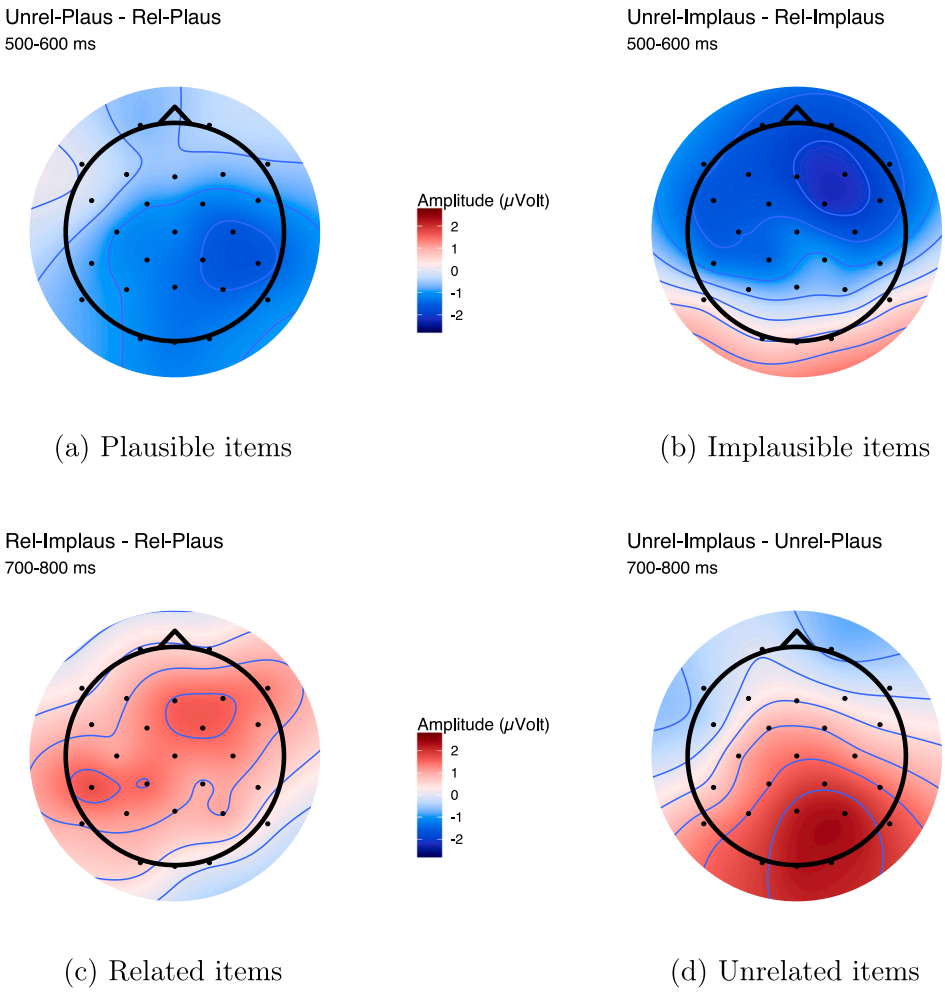



**Fig. C.7.** Topographic maps of the N400 effects of Association for Plausible and Implausible items (top panel) and the P600 effects of Plausibility for Related and Unrelated items (bottom panel).

## Data availability

Data and analysis code are available at https://osf.io/wgjxh/.

## References


Aurnhammer, C., Delogu, F., Brouwer, H., & Crocker, M. W. (2023). The P600 as a continuous index of integration effort. *Psychophysiology*, *60*(9), Article e14302.

Aurnhammer, C., Delogu, F., Schulz, M., Brouwer, H., & Crocker, M. W. (2021). Retrieval (N400) and integration (P600) in expectation-based comprehension. *Plos One*, *16*(9), Article e0257430.

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

Baroni, M., & Kilgarriff, A. (2006). Large linguistically-processed web corpora for multiple languages. In *Demonstrations* (pp. 87–90).

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on semantic P600 effects in language comphrension. *Brain Research Reviews*, *59*, 55–73.

Brothers, T., Wlotko, E. W., Warnke, L., & Kuperberg, G. R. (2020). Going the extra mile: Effects of discourse context on two late positivities during language comprehension. *Neurobiology of Language*, *1*(1), 135–160.

Brouwer, H., & Crocker, M. W. (2017). On the proper treatment of the N400 and the P600 in language comprehension. *Frontiers in Psychology*, *8*, 1327.

Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. J. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science*, *41*, 75–83.

Brouwer, H., Delogu, F., & Crocker, M. W. (2021). Splitting event-related potentials: Modeling latent components using regression-based waveform estimation. *European Journal of Neuroscience*, *53*(4), 974–995.

Brouwer, H., Delogu, F., Venhuizen, N. J., & Crocker, M. W. (2021). Neurobehavioral correlates of Surprisal in language comprehension: A neurocomputational model. *Frontiers in Psychology*, *12*, Article 615538.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143.

Brouwer, H., & Hoeks, J. C. J. (2013). A time and place for language comprehension: Mapping the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience*, *7*, 758.

Brown, C. M., & Hagoort, P. (1993). The processing nature of the N400: evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*, 34–44.

Brown, C. M., & Hagoort, P. (1999). On the electrophysiology of language comprehension: Implications for the human language system. In M. Crocker, P. Pickering, & C. Clifton (Eds.), *Architectures and mechanisms for language processing* (pp. 213–237). Cambridge University Press.

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *NeuroReport*, *18*(17), 1851–1854.

Delogu, F., Brouwer, H., & Crocker, M. W. (2019). Event-related potentials index lexical retrieval (N400) and integration (P600) during language comprehension. *Brain and Cognition*, *135*, Article 103569.

Delogu, F., Brouwer, B., & Crocker, M. W. (2021). When components collide: Spatiotemporal overlap of the N400 and P600 in language comprehension. *Brain Research*, *1766*, 147514.

DeLong, K. A., & Kutas, M. (2020). Comprehending surprising sentences: sensitivity of post-N400 positivities to contextual congruity and semantic relatedness. *Language, Cognition and Neuroscience*, *35*(8), 1044–1063.

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

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, *25*(2), 149–188.

van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18*(7), 1181–1197.

van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicted by semantic anomalies. *Cognitive Brain Research*, *22*, 241–255.

Hoeks, J. C. J., & Brouwer, H. (2014). Electrophysiological research on conversation and discourse processing. In T. Holtgraves (Ed.), *Oxford university handbook of language and social cognition* (pp. 365–386). Cambridge University Press.

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, *19*(1), 59–73.

Kim, A., McKnight, S. M., & Miyake, A. (2024). How variable are the classic ERP effects during sentence processing? A systematic resampling analysis of the N400 and P600 effects. *Cortex*, *177*, 130–149.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*, 205–225.

Klingvall, E., & Heinat, F. (2024). Lexical cues and discourse integration: An ERP study of the N400 and P600 components. *Cortex*, *178*, 91–103.

Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. (2003). Stucture and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, *85*, 1–36.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49.

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, *32*(1), 12–35.

Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*, *100*, 223–238.

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*(1), 117–129.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

Leckey, M., & Federmeier, K. D. (2020). The P3b and P600(s): Positive contributions to language comprehension. *Psychophysiology*, *57*(7), Article e13351.

Luck, S. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Nieuwland, M. S., Barr, D. J., Bartolazzi, F., Busch-Moreno, S., Darley, E., Ferguson, H. J., Fu, X., heyselaar, E., Huettig, F., husband, E. M., Ito, A., Kazanina, N., Kogan, V., Kohút, Z., Kulakova, E., Mézière, D., Politzer-Ahles, S., Rousselet, G., Rueschemeyer, S. A., ...., Segaert, K. et al. (2019). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society of London, B 375*, Article 20180522.

Nieuwland, M. S., & van Berkum, J. J. A. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, *24*(3), 691–701.

Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience*, *23*, 277–293.

Ryskin, R., Stearns, L., Bergen, L., Eddy, M., Fedorenko, E., & Gibson, E. (2021). An ERP index of real-time error correction within a noisy-channel framework of human communication. *Neuropsychologia*, *158*, Article 107855.

Schacht, A., Sommer, W., Shmuilovich, O., Martíenz, P. C., & Martín-Loeches, M. (2014). Differential task effects on N400 and P600 elicited by semantic and syntactic violations. *PloS One*, *9*(3), Article e91226.

Smith, N. J., & Kutas, M. (2015a). Regression-based estimation of ERP waveforms: I. The ERP framework. *Psychophysiology*, *52*, 157–168.

Smith, N. J., & Kutas, M. (2015b). Regression-based estimation of ERP waveforms: II. Nonlinear effects, overlap correction, and practical considerations. *Psychophysiology*, *52*, 169–181.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(1), 176–190.

Venhuizen, N. J., & Brouwer, H. (2025). Referential retrieval and integration in language comprehension: An electrophysiological perspective. *Psychological Review*.

Wang, W., Perfetti, C., & Helder, A. (2024). Getting to the situation: The N400 can indicate meaning integration beyond word priming. *Language, Cognition and Neuroscience*, 1–18.