

# The P600 as a Continuous Index of Integration Effort

Christoph Aurnhammer      Francesca Delogu  
Harm Brouwer      Matthew W. Crocker

December 8, 2022

## Abstract

The integration of word meaning into an unfolding utterance representation is a core operation of incremental language comprehension. There is considerable debate, however, as to which component of the ERP signal – the N400 or the P600 – directly reflects integrative processes, with far reaching consequences for the temporal organisation and architecture of the comprehension system. Multi-stream models maintaining the N400 as integration crucially rely on the availability of a semantically attractive plausible alternative interpretation to account for the absence of an N400 effect in response to certain semantic anomalies, as reported in previous studies. A more recent hypothesis posits the P600 as an index of integration, further predicting that its amplitude varies continuously with integrative effort.

Here we directly test these competing hypotheses using a context manipulation design, in which a semantically attractive alternative is either available or not, and target word plausibility is varied across three levels. An initial self-paced reading study revealed graded reading times for plausibility, suggesting differential integration difficulty. A subsequent ERP study showed no N400 differences across conditions, and that P600 amplitude is graded for plausibility. These findings are inconsistent with the interpretation of the N400 as an index of integration, as no N400 effect emerged even in the absence of a semantically attractive alternative. By contrast, the link between plausibility, reading times, and P600 amplitude supports the view that the P600 is a continuous index of integration effort. More generally, our results support a single-stream architecture, and eschew the need for multi-stream accounts.

## 1 Introduction

In electrophysiological studies of language comprehension, the two most salient components of the Event-Related brain Potentials (ERP) signal are the N400 and the P600. It is still under debate, however, which of these two components indexes semantic integration – the core operation of integrating incoming

32 word meaning into the unfolding utterance representation – during **online** lan-  
33 guage comprehension. Traditionally, semantic integration has been attributed  
34 to the **N400** component (Brown & Hagoort, 1993, 2000; Hagoort, Hald, Basti-  
35 aansen, & Petersson, 2004), such that its amplitude is continuously related to  
36 integration effort, a mapping that underpins several contemporary neurocom-  
37 putational models of comprehension (see Eddine, Brothers, & Kuperberg, 2022,  
38 in prep.). The **P600**, on the other hand, has traditionally been discussed in  
39 relation to syntactic and structural processing (Hagoort, Brown, & Groothusen,  
40 1993; Osterhout & Holcomb, 1992). This linkage of the N400 to semantic inte-  
41 gration and the P600 to structural processing is challenged, however, by studies  
42 employing semantic role violations, such as “the hearty meal was **devouring /**  
43 **devoured**” (Kim & Osterhout, 2005, see also Hoeks, Stowe, & Doedens, 2004;  
44 Kolk, Chwilla, van Herten, & Oor, 2003; Kuperberg, 2007; Kuperberg et al.,  
45 2003; van Herten, Chwilla, & Kolk, 2006; van Herten, Kolk, & Chwilla, 2005),  
46 which lead to P600 rather than N400 effects relative to baseline. To recon-  
47 cile these “semantic P600” findings with the traditional functional roles of the  
48 N400 and the P600, multi-stream models have been proposed which postulate  
49 distinct cognitive mechanisms that trigger either an N400 or a P600, but typi-  
50 cally not both (Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer, Fitz, &  
51 Hoeks, 2012; Kuperberg, 2007, for reviews). Motivated by several limitations of  
52 these models, Retrieval-Integration (RI) theory (Brouwer, Crocker, Venhuizen,  
53 & Hoeks, 2017; Brouwer et al., 2012) offers an alternative, single-stream ac-  
54 count which explains these findings, following an interpretation of the N400  
55 as lexical retrieval (Kutas & Federmeier, 2000, 2011; Lau, Almeida, Hines, &  
56 Poeppel, 2009; Lau, Phillips, & Poeppel, 2008; van Berkum, 2009, 2010) and  
57 reinterpreting the P600 as a continuous index of integration effort.

58 Under multi-stream accounts, the absence of the N400 in the above studies  
59 is explained by positing that an independent semantic processing stream – ef-  
60 fectively a plausibility heuristic that is agnostic to surface structure – fails to  
61 detect the anomaly when a semantically plausible interpretation can be con-  
62 structed from the content words encountered thus far. While different multi-  
63 stream models assign different labels to this heuristic semantic processing stream  
64 (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kim & Osterhout, 2005; Kos,  
65 Vosse, Van Den Brink, & Hagoort, 2010; Kuperberg, 2007; van Herten et al.,  
66 2005), they share the idea that this stream is driven by semantic attraction, i.e.  
67 the availability of a semantically attractive alternative interpretation that goes  
68 against surface structure but that is semantically plausible; e.g., “the hearty



69 meal was “devoured” when reading “devouring” (see Rabovsky, Hansen, & McClelland, 2018; Ryskin et al., 2021 for more recent instantiations of a similar  
70 line of reasoning).

72 A second, algorithmically-driven processing stream, adhering to morphological  
73 and syntactic constraints, does detect the anomalous interpretation constructed by the semantic stream, triggering a P600, thought to reflect a (structural)  
74 revision of the unfolding interpretation. Again, different multi-stream models apply different labels to this algorithmic-processing stream (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kim & Osterhout, 2005; Kos et al., 2010; Kuperberg, 2007; van Herten et al., 2005), but the shared intuition is that a conflict between the heuristic semantic stream and an algorithmic-driven stream  
75 yields a constellation that requires a revision of the unfolding interpretation (see  
76 Rabovsky & McClelland, 2020; Ryskin et al., 2021, for recent incarnations of  
77 this idea). Crucially, the absence of an N400 together with the presence of a  
78 P600 effect compared to baseline depends on the availability of a repairable  
79 anomaly, for instance because thematic roles can be reversed. Conversely, if no  
80 such alternative is present, multi-stream models predict an N400 effect indexing  
81 increased integrative effort, but no P600 effect relative to the baseline condition.

82 By contrast, RI theory postulates a single stream account in which the N400  
83 is taken to index lexical retrieval (Kutas & Federmeier, 2000, 2011; Lau et al.,  
84 2009, 2008; van Berkum, 2009, 2010), and the P600 is taken to index semantic  
85 integration effort (Brouwer et al., 2017, 2012). The lexical retrieval of word  
86 meaning from long term memory is facilitated – and N400 amplitude attenuated  
87 – when the meaning of an incoming word is lexically or contextually primed. In  
88 the case of “the hearty meal was devouring / devoured”, the word meanings of  
89 “devouring” and “devoured” are equally primed by the “the hearty meal”, therefore  
90 producing no N400 effect between the violating and non-violating verbs.  
91 Semantic integration of the retrieved meaning of the current word with the un-  
92 folding utterance interpretation in working memory is, however, more difficult  
93 for the implausible case in which the “meal” is the agent of the “devouring”  
94 action compared to the plausible case in which it is the patient. This integra-  
95 tion difficulty is reflected in the P600 effect obtained for this contrast. The  
96 absence of an N400 together with the presence of a P600 effect relative to base-  
97 line is thus in line with Retrieval-Integration theory, and eschews the need to  
98 invoke separate processing streams. Further, this interpretation is strengthened  
99 by “semantic P600s” that were not elicited by role-reversals, but that rather  
100 reflect general semantic and pragmatic processing effort involved establishing a  
101  
102  
103  
104  
105

106 coherent representation (Burkhardt, 2006, 2007; Cohn & Kutas, 2015; Delogu,  
107 Brouwer, & Crocker, 2019; Dimitrova, Stowe, Redeker, & Hoeks, 2012; Hoeks,  
108 Stowe, Hendriks, & Brouwer, 2013; Regel, Gunter, & Friederici, 2010; Schu-  
109 macher, 2011; Spotorno, Cheylus, Henst, & Noveck, 2013; Xu & Zhou, 2016).  
110 Indeed, on RI theory, every word is predicted to modulate P600 amplitude, re-  
111 flecting the relative ease with which a word's meaning can be integrated into the  
112 current utterance meaning representation. Importantly, it follows that the P600  
113 should not be a binary response to violating stimuli, but rather, P600 ampli-  
114 tude should be sensitive to integration difficulty on a continuous scale (Brouwer  
115 et al., 2012), and indeed linked to Comprehension-centric Surprisal (Brouwer,  
116 Delogu, & Crocker, 2021). Preliminary evidence for this graded link has been  
117 presented in a post-hoc analysis by Aurnhammer, Delogu, Schulz, Brouwer, and  
118 Crocker (2021), who demonstrated a graded response of both the N400 and the  
119 P600 for congruous sentences that varied in target word expectancy.

120 Multi-stream models and RI theory thus critically differ in that they at-  
121 tribute semantic integration to the N400 and the P600, respectively, and hence  
122 make different predictions for which component should reveal an integration  
123 effect in absence of a semantically attractive alternative, relative to a congru-  
124 ous baseline condition. Here, we present an experimental design that directly  
125 tests the predictions of multi-stream models against those of RI theory. To this  
126 end, we build on the design by Nieuwland and van Berkum (2005) in which a  
127 context paragraph is followed by a critical region including either a plausible  
128 (coherent: “the woman told the tourist”) or an implausible (incoherent: “the  
129 woman told the suitcase”) target word (Table 1). Crucially, both target words,  
130 “tourist” and “suitcase”, are mentioned several times in the preceding context  
131 paragraph. Stimuli were presented in spoken form and without a task. The  
132 contrast of the implausible (incoherent) “suitcase” to the plausible (incoherent)  
133 “tourist” elicited a broadly distributed P600 effect, but no N400 effect.

134 This result seems inconsistent with multi-stream accounts. That is, when  
135 encountering the implausible target word “suitcase”, there is no *locally* available  
136 semantically attractive alternative – e.g., through sentence-internal permutation  
137 of thematic roles and/or morphological inflection – that would yield a plausi-  
138 ble interpretation of the sentence. As a result, multi-stream models predict  
139 an N400 effect, reflecting the difficulty in arriving at a semantically plausible  
140 analysis when compared to a plausible sentence, but no P600 effect, as there  
141 is no disagreement between the independent semantic stream and the algo-  
142 rithmic stream (see Brouwer et al., 2012 for discussion; a schematic multi-stream

### *Introduction*

A tourist wanted to bring his huge suitcase onto the airplane. However, because the suitcase was so heavy, the woman behind the check-in counter decided to charge the tourist extra. In response, the tourist opened his suitcase and threw some stuff out. So now, the suitcase of the resourceful tourist weighed less than the maximum twenty kilos.

#### *Coherent continuation*

Next, the woman told the tourist that she thought he looked really trendy. The tourist grabbed the woman's hand and eagerly asked her for a date. But the woman reprimanded the tourist for being pushy and told him to just get on the plane right away.

#### *Incoherent continuation*

Next, the woman told the suitcase that she thought he looked really trendy. The suitcase grabbed the woman's hand and eagerly asked her for a date. But the woman reprimanded the suitcase for being pushy and told him to just get on the plane right away.

Table 1: Experimental stimulus from the design of Nieuwland and van Berkum (2005), translated from Dutch.

143 analysis is given in Figure 1, left). It has been argued, however, that a semanti-  
144 cally attractive alternative may be globally available in the larger discourse (see  
145 Bornkessel-Schlesewsky & Schlesewsky, 2008; Kuperberg, 2007, for discussion).  
146 That is, as both “tourist” and “suitcase” are salient entities in the discourse,  
147 which have been mentioned numerous times, the interpretation of the coherent  
148 condition (“the woman told the tourist”) may actually be a strong attractor in  
149 the incongruent condition; in other words, the salience of the plausible noun  
150 phrase “the tourist” may *distract* the system away from the actual noun phrase  
151 “the suitcase”. If this is the case, a multi-stream account of this result would  
152 entail the independent semantic stream encountering no difficulty in producing  
153 a plausible analysis, which should lead to no N400 modulation, thereby yielding  
154 a conflict with the algorithmic processing stream (which arrives at the analy-  
155 sis “the woman told the suitcase”), thereby triggering a P600 effect relative to  
156 baseline (see Figure 1, right).

157 Retrieval-Integration theory, on the other hand, attributes the absence of  
158 an N400 effect to facilitated retrieval. That is, the lexical repetition of both  
159 the congruent and incongruent target words leads to maximal priming of their  
160 meaning. Indeed, in line with this interpretation, the N400 effect resurfaced,  
161 for similar stimuli presented in story-initial position, i.e., without any preceding  
162 context mentioning the target words (see Figure 4 in Nieuwland & van Berkum,

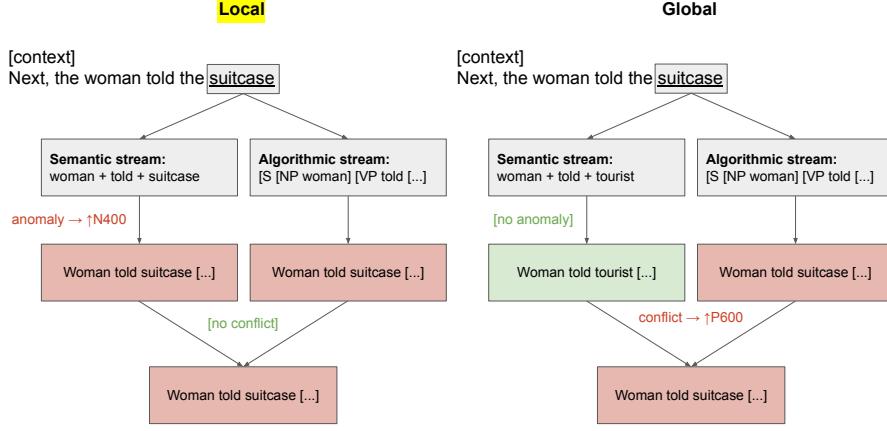


Figure 1: Schematic overview on multi-stream explanations assuming either a *local* or a *global* revision mechanism.

2005), due to the absence of equal priming for “suitcase” and “tourist”.<sup>1</sup> The presence of a P600 effect, in turn, reflects the increased difficulty in integrating “suitcase” versus “tourist” in “the woman told [...], as the former yields an interpretation that goes against world knowledge. If we accept the independent processing streams of multi-stream models to be able to compute a *globally* available semantically attractive alternative interpretation, then multi-stream models and RI theory make the same N400 and P600 predictions, and both account for the Nieuwland and van Berkum (2005) data. Crucially, however, if no such alternative interpretation is available, the accounts make diverging predictions: multi-stream models predict an N400 and no P600, and RI theory predicts a P600 effect relative to baseline. While previous studies observing semantic P600 effects typically employed binary designs, RI theory makes the specific prediction that P600 amplitude should be a function of graded integration difficulty (see Aurnhammer et al., 2021, for preliminary support). To test these diverging predictions, we here present an adapted version of the Nieuwland and van Berkum (2005) design.

Specifically, the new design implements several manipulations (see Table 2). First, we created a baseline condition, in which the target word is expected

<sup>1</sup>Visual inspection suggests that this N400 effect co-occurs with an increase in P600 amplitude.



and plausible and no processing difficulties should ensue (**Condition A**). In order to test the prediction of multi-stream models that it is the availability of a semantically attractive alternative that explains the absence of an N400 effect and the presence of a P600 effect, we constructed one condition such that an alternative is made *globally* available by a distractor word in the context (**Condition B**). In another condition, no such alternative is available (**Condition C**) and we compare both conditions to the unmanipulated baseline (**Condition A**). Further, to test for the gradedness of integration difficulty, the target word in Condition B has intermediate plausibility, in that it renders the interpretation semantically unlikely yet possible, while Condition C is implausible, yielding a semantic anomaly (see Figure 4 for more examples). Finally, to maximise comparability of target word processing across conditions, our design employs a context rather than a target manipulation design and we harness lexical repetition to maximally and equally prime the target words in the three conditions.

Multi-stream models predict a P600 and no N400 effect for Condition B relative to Condition A. This is because the anomaly is repairable by replacing the anomalous interpretation resulting from the observed word with the *globally* available alternative interpretation that derives from the distractor word, similar to the original study. In Condition C however, no such alternative interpretation is licensed by the context and hence multi-stream models predict an N400 effect, and critically, no P600 effect relative to the baseline condition. RI theory, on the other hand, predicts that no N400 differences should be produced across conditions due to the lexical repetition of the target word in the context paragraph, maximally facilitating lexical retrieval of its meaning. Under the hypothesis that P600 amplitude continuously indexes the effort of integrating word meaning with the utterance meaning representation constructed so far, the P600 is predicted to be graded for plausibility with increasing amplitude for conditions  $A < B < C$ . In sum, while multi-stream models predict a P600 effect for Condition B, and an N400 effect for Condition C relative to the baseline Condition A, RI theory predicts the absence of N400 effects, and graded P600 amplitude differences across conditions (see Table 3).

On the assumption that reading times provide an index of overall word-by-word processing difficulty, we first collected self-paced reading time data for our novel design. We expect that reading times should be graded for target word plausibility, reflecting graded integration difficulty. Subsequently, we recorded Event-Related Potentials for the same stimuli, allowing for a direct comparison between behavioral and neurophysiological indices of integrative processing

*Context*

Ein Tourist wollte seinen riesigen **Koffer** mit in das Flugzeug nehmen. Der **Koffer** war allerdings so schwer, dass die Dame am Check-in entschied, dem Touristen eine extra Gebühr zu berechnen. Daraufhin öffnete der Tourist seinen **Koffer** und warf einige Sachen hinaus. Somit wog der **Koffer** des einfallsreichen Touristen weniger als das Maximum von 30 Kilogramm.

*A tourist wanted to take his huge suitcase onto the airplane. The suitcase was however so heavy that the woman at the check-in decided to charge the tourist an extra fee. After that, the tourist opened his suitcase and threw several things out. Now, the suitcase of the ingenious tourist weighed less than the maximum of 30 kilograms.*

*Condition A: Plausible, no attraction*

Dann verabschiedete die Dame den Touristen und danach ging er zum Gate.  
*Then dismissed the lady the tourist and afterwards he went to the gate.*

*Condition B: Less plausible, attraction*

Dann wog die Dame den Touristen und danach ging er zum Gate.  
*Then weighed the lady the tourist and afterwards he went to the gate.*

*Condition C: Implausible, no attraction*

Dann unterschrieb die Dame den Touristen und danach ging er zum Gate.  
*Then signed the lady the tourist and afterwards he went to the gate.*

Table 2: Experimental design of the present study. German word order is preserved for the English transliterations of the final sentences. Target words are underlined and distractor words are highlighted in boldface.

	Multi-stream		Retrieval-Integration	
	N400	P600	N400	P600
A: Plausible, no attraction	-	-	-	-
B: Less plausible, attraction	-	+	-	+
C: Implausible, no attraction	+	-	-	++



Table 3: Predictions of multi-stream models and Retrieval-Integration theory for the N400 and the P600 component in the current design.

<sup>218</sup> effort (see Brouwer, Delogu, Venhuizen, & Crocker, 2021, for discussion).

<sup>219</sup> **2 Experiment 1: Self-paced Reading**

<sup>220</sup> **2.1 Method**

<sup>221</sup> Code and data required to reproduce the analyses will be made publicly avail-  
<sup>222</sup> able upon acceptance. All studies were conducted with ethics approval of the  
<sup>223</sup> Deutsche Gesellschaft für Sprachwissenschaft (DGfS).

<sup>224</sup> **2.1.1 Materials**

<sup>225</sup> The materials were optimised to be used in the same form in the self-paced  
<sup>226</sup> reading study and the EEG study (see Appendix A for the full list of German  
<sup>227</sup> stimuli). In the creation of the stimuli, we translated and adapted items from  
<sup>228</sup> Nieuwland and van Berkum (2005) where possible, and otherwise developed  
<sup>229</sup> new items. In total, we developed 96 items for which we changed the original  
<sup>230</sup> target manipulation to a context manipulation design. Employing a context  
<sup>231</sup> manipulation design in which the target word is the same across conditions is  
<sup>232</sup> intended to reduce effects due to differences in word length, frequency, etc. Each  
<sup>233</sup> item had the same context paragraph for each condition. The context paragraph  
<sup>234</sup> repeatedly mentioned both the target word as well as a distractor word. The  
<sup>235</sup> target word and the distractor word were mentioned the same amount of times  
<sup>236</sup> within item (three or four times). Presenting the target word several times  
<sup>237</sup> in the context paragraph should maximally prime the target word's meaning,  
<sup>238</sup> when presented in target position. Under RI theory, we thus expect no N400  
<sup>239</sup> (retrieval) effect across conditions (cf. Brouwer & Crocker, 2017; Brouwer et al.,  
<sup>240</sup> 2012). Which of the two words – target or distractor – was last mentioned in  
<sup>241</sup> the context was approximately balanced across items.



<sup>242</sup> The context paragraph was followed by a manipulated final sentence. Conditions differed only in the main verb of the final sentence, rendering the target word of the sentence – i.e. the direct object – plausible (Condition A, “the lady dismissed the tourist”), less plausible (Condition B, “the lady weighed the tourist”), or implausible (Condition C, “the lady signed the tourist”). Taken together, this allows us to assess whether plausibility results in graded effects on both RTs and the P600. Additionally, the distractor word, which is never presented in target position, was either expected (Condition B, “the lady weighed” attracting “suitcase”), or unexpected (Condition A and C), allowing us to investigate whether the presence of a semantically attractive alternative interpretation modulates the presence of P600 (Condition B; semantic attraction) or



253 N400 effects (Condition C; no semantic attraction) in the ERP experiment. 254 The final sentence of each item ended with an additional clause following the target word (255 “[...] **and** afterwards he went to the gate”), which avoids placement of 256 the target in sentence-final position and allows us 257 to capture spillover effects in reading times. Figure 4 shows four more transliterated items.

#### *Item 2*

A teacher saw an old world map in the showcase of an antique shop. Such an authentic artefact appeared suitable for his classroom and he approached the **saleswoman**...

*A*: Then bought the teacher the map ...

*B*: Then kissed the teacher the map ...

*C*: Then filled the teacher the map ...

#### *Item 4*

While building a table, a **carpenter** broke his nice hammer into pieces...

*A*: Then took the apprentice the hammer ...

*B*: Then sneered-at the apprentice the hammer ...

*C*: Then ate the apprentice the hammer...

#### *Item 11*

In a foreign city, a vacationer booked a guided tour. The **guide** was happy that the vacationer was interested and gifted him a flyer...

*A*: After the tour folded the vacationer the flyer ...

*B*: After the tour commended the vacationer the flyer ...

*C*: After the tour cooked the vacationer the flyer ...

#### *Item 18*

A young lady wanted to have a **jewel** evaluated by a jeweller...

*A*: Delighted remunerated the lady the jeweller ...

*B*: Delighted marveled-at the lady the jeweller ...

*C*: Delighted seasoned the lady the jeweller ...

Table 4: Four example items, transliterated from German. Target words are underlined, distractor words are highlighted in boldface.

258 **Cloze Norming** We collected cloze probabilities to validate the differential 259 expectancy of both the target and distractor words across conditions. Sentence 260 completions were collected in a web-based experiment using the software PCIbex 261 (Zehr & Schwarz, 2018), which we also used for all other web-based norming 262 studies and experiments reported here. 263 We did not use filler items, since the materials up to the target word do not contain any anomalies. Participants



264 were presented with the entire context paragraph and the final sentence up to  
265 – but not including – the determiner of the target word. That is, we did not  
266 provide a determiner as the grammatical gender of German would constrain the  
267 set of possible completions. Cloze probabilities were obtained in two rounds,  
268 both optimising the contrast of high expectation for the target (Condition A) or  
269 the distractor word (Condition B). Sentence contexts for implausible words were  
270 created such that they do not raise strong expectations for any specific word  
271 (Condition C). In total, we collected responses from 90 participants, who were  
272 recruited through Prolific Academic Ltd. and each was paid £7.50. We selected  
273 the 60 best items based on the results of the cloze task. Alternative cloze prob-  
274 abilities for any other word in Condition C were kept below 0.27 (mean = 0.20;  
275 SD = 0.07). The resulting cloze probabilities for the target and distractor word  
276 across the three conditions are presented in Table 5 and Figure 2 (left). Target  
277 word cloze probability is high in Condition A, indicating high expectancy of the  
278 target word in the baseline condition, which should therefore induce only low  
279 integrative effort. In Condition B, participants actively produced the distractor  
280 word rather than the target word, indicating that the distractor word indeed  
281 makes a semantically attractive alternative interpretation available in this con-  
282 dition. In Condition C, expectancy of both the target word and the distractor  
283 word were low. The latter suggests that the alternative interpretation avail-  
284 able for Condition B is removed in Condition C. In sum, the cloze probabilities  
285 suggest that the availability of the semantically attractive alternative interpre-  
286 tation has been manipulated successfully (Condition A: baseline; Condition B:  
287 semantic attraction; Condition C: no semantic attraction). We turn to a second  
288 norming study in which we collect plausibility ratings to discern whether the  
289 target words of conditions B and C – which were similarly unexpected – indeed  
290 differ in their plausibility.

291 **Plausibility** In a second norming study, we collected plausibility ratings for  
292 the target and distractor words on a seven-point Likert scale, 7 indicating “very  
293 plausible” and 1 indicating “not plausible”. In total, 60 participants were re-  
294 cruited through Prolific Academic Ltd., and each was paid £7.50. For the rating  
295 task, the final sentence was presented in one paragraph together with the con-  
296 text material, with the aim to ensure reading of the entire paragraph and not  
297 only the final sentence. Participants were instructed to rate the plausibility  
298 of the final sentence in light of the context. We excluded the final sentence  
299 continuation (“and afterwards he went to the gate”) to maximise rating the



		Cloze			Plausibility		
	Cond.	Mean	SD	Range	Mean	SD	Range
Target	A	0.80	0.20	0.33-1.00	5.84	0.93	3.60-7.00
	B	0.09	0.11	0.00-0.40	3.69	1.33	1.50-6.30
	C	0.02	0.04	0.00-0.20	1.42	0.33	1.00-2.40
Distractor	A	0.05	0.90	0.00-0.33	2.53	1.34	1.10-6.30
	B	0.78	0.17	0.33-1.00	5.94	1.05	2.40-7.00
	C	0.03	0.06	0.00-0.20	1.66	0.69	1.00-4.80

Table 5: Averages, standard deviations, and ranges for the results of two norming studies that collected cloze probabilities and seven-point scale plausibility ratings for the target and the distractor word.

target word rather than another part of the final sentence. During the rating task, there were 10 items with attention checks which presented mid-paragraph instruction to rate this trial with a given number (either 1 or 7). On average, participants completed 98 % of attention checks successfully (mean = 98.19 %; SD = 4.09; range 83.33 % - 100.00 %). The resulting plausibility ratings are reported in Table 5 and Figure 2 (right). Target word plausibility is stepped across conditions (A > B > C), which should result in a similarly graded effect of integration effort of the target in the three contexts. Distractor word plausibility is high in Condition B while in Condition A and C, distractor word plausibility is low, again supporting the availability of a semantically attractive alternative interpretation in Condition B.

Correlations between target and distractor word cloze probability and plausibility are reported in Table 6. Our analyses will focus on target word plausibility to investigate graded effects of plausibility and on distractor cloze to investigate additional effects of semantic attraction. As the correlations show, these predictors are effectively independent ( $r = 0.01$ ).

		Cloze		Plausibility	
		Target	Distractor	Target	Distractor
Cloze	Target	1.00	-0.40	0.79	-0.24
	Distractor	-0.40	1.00	0.01	0.88
Plausibility	Target	0.79	0.01	1.00	0.22
	Distractor	-0.24	0.88	0.22	1.00

Table 6: Correlations between cloze probabilities and plausibility ratings of the target and distractor words.

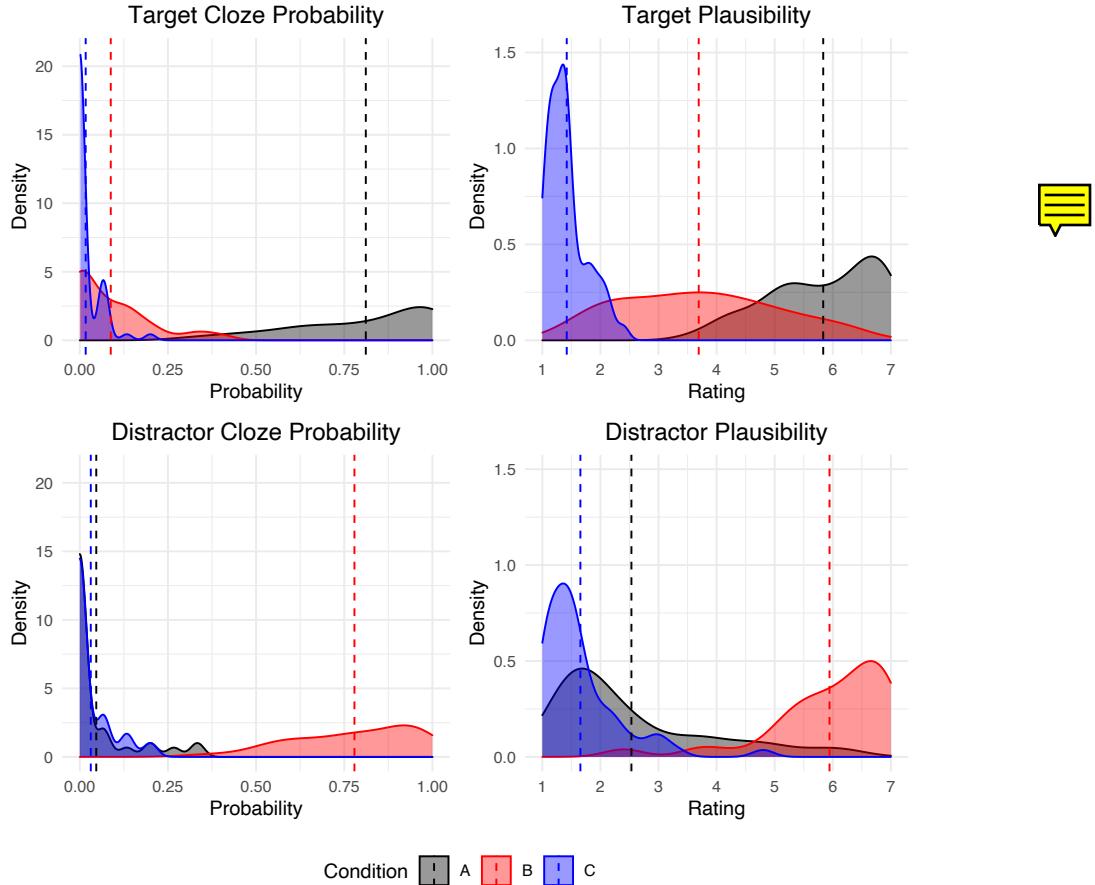


Figure 2: Densities for the results of two norming studies that collected cloze probabilities and seven-point scale plausibility ratings for the target and the distractor words.

### 316 2.1.2 Participants

317 Forty-three participants were recruited through Prolific Academic Ltd., to take  
 318 part in the web-based self-paced reading experiment. One participant was ex-  
 319 cluded due to inattentive reading, as shown by low accuracy on the task (60%  
 320 correct; see below for specifics of the task). The remaining 42 participants  
 321 (mean age 24.43; SD 3.7; age range 18-32; 15 male, 27 female) were all native  
 322 speakers of German (2 early bilinguals) and had not indicated any language  
 323 related disorder or literacy difficulty. They did not participate in any other  
 324 studies reported in this article. All participants gave their consent by agreeing

325 to a consent form and were paid £7.50 for their participation.

326 **2.1.3 Procedure**

327 We conducted the self-paced reading experiment as a web-based study. On  
328 each trial, participants were prompted to press the Enter-key to start, after  
329 which they were presented with a context paragraph. Upon pressing the Enter-  
330 key again, a hash sign was presented centrally, indicating the position of the  
331 words of the final sentence. From here on, participants pressed the Space-bar to  
332 proceed to the next word, each presented centrally. After three practice items,  
333 stimuli were presented in three blocks with 35 items each, summing to a total  
334 of 105 items, 45 of which were fillers. For half of the participants, the blocks  
335 and the items within them were presented in reverse order. On 46% of trials –  
336 half of the experimental trials and on two fifth of the fillers – participants were  
337 presented with a comprehension question to which they had to answer using  
338 either Yes or No (mapped to the *D* and *K* keys). Comprehension questions had  
339 Yes and No as correct answer on 50% of the questions and they could concern  
340 the context paragraph or the final sentence, within which they could focus on  
341 the manipulated region or the final sentence completion. To encourage attentive  
342 reading, we provided coarse feedback on participants' response accuracy after  
343 the practice session and after each block. Participants were encouraged to take  
344 a short break between blocks.

345 **2.1.4 Analysis**

346 We excluded trials if reading time on any critical region was lower than 50 ms or  
347 higher than 2500 ms and if reaction time on the task (if there was one on that  
348 trial) was lower than 50 ms or higher than 10,000 ms. Based on these criteria,  
349 47 out of 2520 trials were excluded (1.87%). All results and analyses reported  
350 below are computed after exclusion.



351 Log-transformed reading times were analysed using a linear mixed effects re-  
352 gression re-estimation technique (cf. Aurnhammer et al., 2021). This technique  
353 fits reading time models separately on each region of interest, allowing to trace  
354 across regions the relative influence and significance of each predictor in the  
355 regression equation as well as the residual error, i.e. the difference between the  
356 observed data and the forward estimates computed by the models. As predictors  
357 of interest, we focus on target word plausibility and distractor cloze probabili-  
358 ty. Plausibility ratings will serve as a continuous predictor to operationalise

359 integration difficulty of the target word. Distractor cloze probability serves as  
 360 a predictor that will explain any additional effort incurred by the availability of  
 361 a semantically attractive alternative interpretation. Random intercepts as well  
 362 as random slopes for each predictor are estimated for both subjects and items.  
 363 The full model specification is

364

$$Y = \beta_0 + S_0 + I_0 + (\beta_1 + S_1 + I_1)Plaus + (\beta_2 + S_2 + I_2)Clozedist + \epsilon \quad (1)$$

365 in which  $\beta_0$  represents the fixed-effect intercept and  $\beta_1$  and  $\beta_2$  refer to the  
 366 fixed-effect coefficients of plausibility and distractor cloze probability.  $S$  and  
 367  $I$  terms represent random intercepts and slopes for subjects and items. The  
 368 unexplained variance in the data is represented by the residual error term  $\epsilon$ . All  
 369 predictors were standardized, centering their average value on zero and expressing  
 370 them on a scale of standard deviations. Standardising predictors additionally  
 371 has the effect that the intercept will equal the mean of the data to which the  
 372 model is fitted. Plausibility was also inverted, as we predict that higher reading  
 373 times ensue for low plausibility ratings. We run separate analyses for the dif-  
 374 ferent regions of interest, and we treat them as separate families of hypotheses  
 375 and hence do not correct for multiple comparisons.

## 376 2.2 Results

### 377 2.2.1 Comprehension Questions

378 Participants answered comprehension questions on half of the experimental  
 379 items. Descriptive metrics for accuracy and reaction times were computed across  
 380 subjects. Average accuracy was 95.6% ( $SD = 5.3$ , range = 80.0% - 100.0%).  
 381 Mean reaction time was 3098 ms ( $SD = 619$ , range = 1907 ms - 4426 ms).  
 382 Accuracies and reaction times per condition are given in Table 7.

Cond.	Accuracy			Reaction Time		
	Mean	SD	Range	Mean	SD	Range
A	96.7%	6.1	80.0% - 100.0%	2982 ms	651	1566 ms - 4278 ms
B	95.0%	8.0	70.0% - 100.0%	3132 ms	605	1986 ms - 4632 ms
C	95.2%	7.1	80.0% - 100.0%	3184 ms	764	2086 ms - 5190 ms

Table 7: SPR task performance. Task performance on the comprehension questions in the self-paced reading experiment. Accuracy and reaction times were computed across subjects.

383    **2.2.2 Reading Times**

384    Figure 3 displays log-transformed reading times, split up per condition, on the  
 385    Pre-critical region (the ambiguous article “den” / “the” of the target word), the  
 386    Critical region (the target word “tourist”), the Spillover region (“and”) and the  
 387    Post-spillover region (“afterwards”). Visual inspection of the data suggest that  
 388    already on the Pre-critical region, Condition C is read slower than Condition A  
 389    and B. On the Spillover region, Condition B and C are slowed down. Lastly, on  
 390    the post-spillover region, reading times appear to pattern with the three levels  
 391    of Conditions A, B, and C.

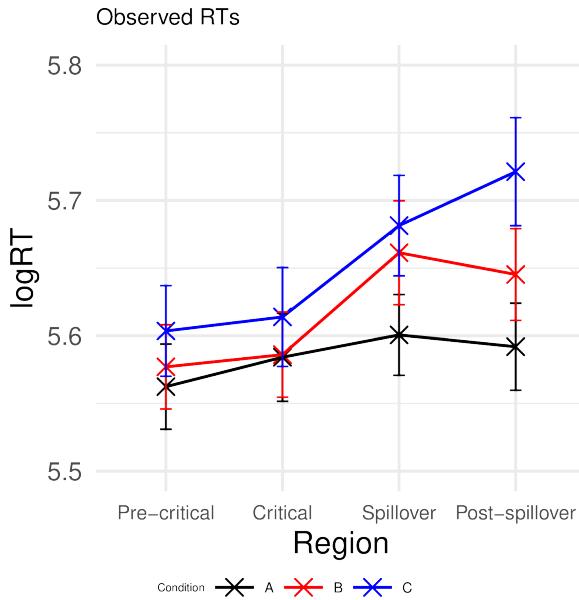


Figure 3: Log reading times per condition on the Pre-critical, Critical, Spillover, and Post-spillover region. Error bars indicate the standard error computed from the per-subject per-condition averages.

392    We modeled the reading times as a function of target word plausibility and  
 393    distractor cloze probability separately on each region. Figure 4 displays the  
 394    estimated reading times from these models as well as the residual error, i.e.  
 395    the difference between the observed and the estimated reading times. Visual  
 396    inspection suggests that the models capture the effects structure in the observed  
 397    data as evidenced by small residual error across regions and conditions.

398    Figure 5 (left) displays model coefficients, added to their intercept, for plau-

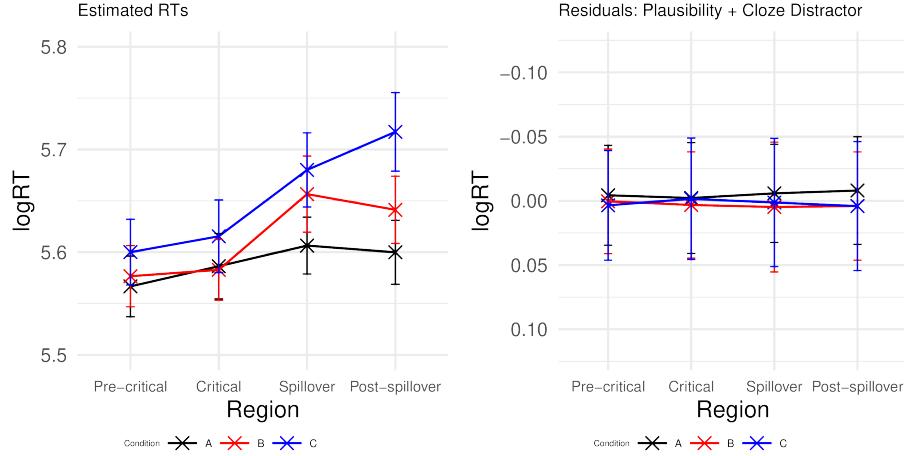


Figure 4: Estimated log-Reading Times (left) and residual error (right) per condition on the Pre-critical, Critical, Spillover, and Post-spillover regions.

399 sibility and distractor cloze probability together with their respective z- and  
 400 p-values (right). The positive coefficients for plausibility indicate that lower  
 401 plausibility predicts slower reading. The coefficient for distractor cloze proba-  
 402 bility is smaller and changes sign moving from the Critical to the Spillover and  
 403 to the Post-spillover region, indicating that this predictor estimates slower or  
 404 faster reading time depending on the region of interest. The z- and p-values  
 405 demonstrate that target word plausibility significantly predicts reading times  
 406 across all regions, interestingly also including the Pre-critical region, while no  
 407 significant contribution of distractor cloze probability was found.

408 Reading in the implausible Condition C is slowed already prior to the target  
 409 word, presumably due to differences in processing of the main verbs preceding  
 410 the targets. This raises the question to what extent reading time differences  
 411 observed on and after the critical word are due to the plausibility of the target  
 412 word itself, rather than due to the different contexts. To answer this question,  
 413 we included the reading time on the Pre-critical region as a predictor into our  
 414 model, allowing the models to capture any pre-critical reading time offsets. We  
 415 only z-scored but did not log-transform the Pre-critical RT predictor, in order  
 416 to avoid identity of the dependent (logRT) and an independent variable (Pre-  
 417 critical RT) on the Pre-critical region. The remaining predictors now explain  
 418 any systematic variability in reading time over and above reading time offsets  
 419 present at the Pre-critical region. The resulting coefficients and z-values indicate

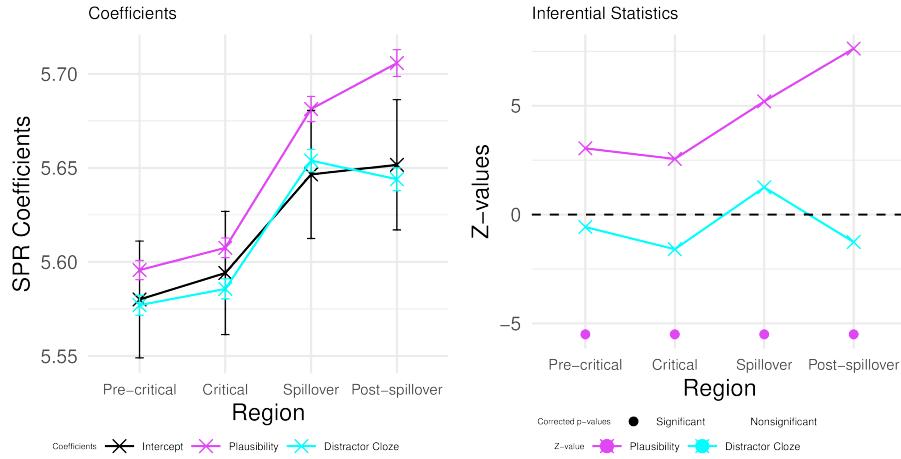


Figure 5: Coefficients (left; added to their intercept), effect sizes (z-values) and p-values (right). Error bars indicate the standard error of the coefficients in the fitted statistical model.

420 that target word plausibility significantly predicts slowed reading time at the  
421 Spillover and Post-Spillover regions, over and above what is explained by Pre-  
422 critical reading time, whereas it is no longer significant on the Pre-critical and  
423 Critical regions. Distractor cloze probability still does not significantly predict  
424 reading times on any region (Figure 6).

### 425 2.3 Discussion

426 The results of the self-paced reading experiment show that reading times scale  
427 gradually with plausibility, indicating that our manipulation of target plausibil-  
428 ity indeed resulted in a graded modulation of integration difficulty. Furthermore,  
429 the regression-based analysis revealed that plausibility is a continuous predictor  
430 of reading time.

431 Based on the traditional surprisal literature (Frank, Otten, Galli, & Vigliocco,  
432 2015; Levy, 2008; Monsalve, Frank, & Vigliocco, 2012), it could be expected that  
433 the same items that show modulations in reading times would also elicit a graded  
434 N400 response. However, the hypothesis that the P600 reflects integration effort  
435 predicts a strong link between this component and late reading time measures  
436 (Brouwer & Crocker, 2017; Brouwer et al., 2012). Empirical evidence in sup-  
437 port of this is provided by Brouwer, Delogu, Venhuizen, and Crocker (2021)  
438 and Aurnhammer et al. (2021) showing that reading time modulations pattern

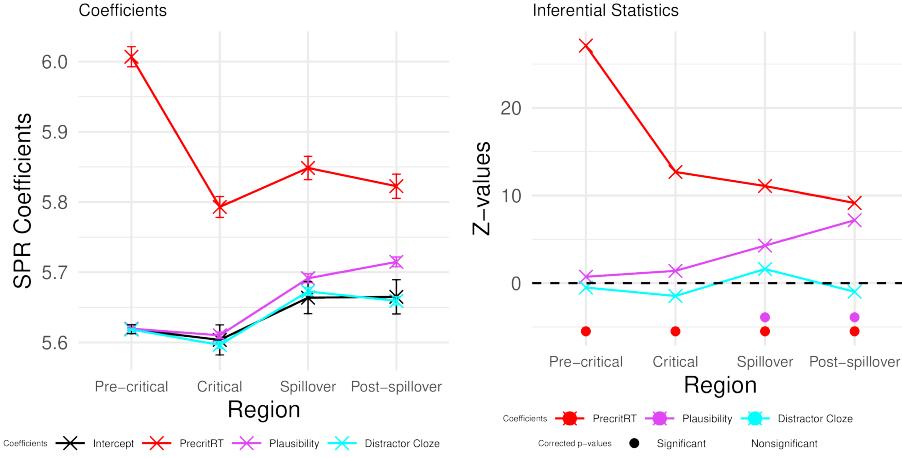


Figure 6: Coefficients (left; added to their intercept), effect sizes (z-values) and p-values (right) from a model including Pre-critical reading time as a predictor. Error bars indicate the standard error of the coefficients in the fitted statistical models.

439 with P600 effects.<sup>2</sup> The obtained reading times thus offer an opportunity to investi-  
440 ginate whether the experimental design will result in a graded N400 or P600 pattern.  
441

442 The current results did not reveal significant reading time modulations due  
443 to distractor cloze probability. Hence, no significant reading time modulation  
444 can be attributed to the presence of a semantically attractive alternative in-  
445 terpretation in Condition B. Our manipulation effectively creates a prediction  
446 disconfirmation, since in the context “Then weighed the lady”, the expected  
447 word “suitcase” is not presented, while “tourist” is provided instead. Previous  
448 research on prediction error cost has not found disconfirmation effects in the  
449 behavioural domain using eye-tracking (Frisson, Harvey, & Staub, 2017; Luke  
450 & Christianson, 2016) or self-paced reading (Rich & Harris, 2021). In a self-  
451 paced reading experiment by Van Berkum, Brown, Zwitserlood, Kooijman, and  
452 Hagoort (2005) a disconfirmation effect was observed - however its timing did  
453 not coincide with the ERP deflection found for the same stimuli. Similarly, lex-  
454 ical decision times did not exhibit facilitation effects for unrelated, unexpected  
455 words in high constraint sentences relative to the same words in low constraint

<sup>2</sup> Additionally, effects of association, which were also reflected in N400 amplitude, modulated reading times on the first Spillover region. As the current design maximally primes the targets across all conditions, no such association-related effects were expected in the current data.

456 sentences (Schwanenflugel & LaCount, 1988). In line with this previous re-  
457 search, our results suggest that reading times may not be sensitive to unfulfilled  
458 expectations. With regard to the comparison of multi-stream models and RI  
459 theory, the absence of a significant contribution of semantic attraction (dis-  
460 tractor cloze probability) in behavioural measures raises the question whether  
461 semantic attraction will modulate the presence of P600 and N400 effects in the  
462 ERP signal.

### 463 3 Experiment 2: Electroencephalography

#### 464 3.1 Method

##### 465 3.1.1 Materials

466 The materials were the same as in the self-paced reading experiment (see Section  
467 2.1.1).

##### 468 3.1.2 Participants

469 We recruited 33 participants at Saarland University to take part in the ex-  
470 periment. Three participants were excluded due to excessive eye-movement  
471 artefacts. The final 30 participants (mean age 25; SD = 3.35; range 18-32; 25  
472 female, 5 male) were right-handed, native speakers of German (six early bilin-  
473 guals) and had normal or corrected-to-normal vision. None reported any form  
474 of color blindness. Participants gave informed, written consent and were paid  
475 25€.

##### 476 3.1.3 Procedure

477 We recorded the electroencephalogram (EEG) while the participants were seated  
478 in an electromagnetically shielded, sound-proof, and dimly lit chamber. The  
479 experiment was presented using E-prime 3 (Schneider, Eschman, & Zuccolotto,  
480 2002). We first presented three practice items, two of which included a com-  
481 prehension question. Practice items varied in their degrees of plausibility. The  
482 practice session was followed by three blocks, each containing 35 items, includ-  
483 ing the same fillers as for the self-paced reading experiment. Participants took  
484 a break between blocks. Items were presented in pseudorandomized order. For  
485 half of the participants, the blocks and the items within them were presented in  
486 reverse order. On each trial, participants used a button-box to start the item

487 and were presented with the entire context paragraph which remained on the  
488 screen until the button was pressed again. Then, a fixation cross appeared in  
489 the center of the screen for 750 ms. After that, the final sentence was presented  
490 using rapid serial visual presentation (RSVP). Each word of the final sentence  
491 was presented centrally for 350 ms with a 150 ms inter-stimulus interval. If  
492 the item contained a comprehension question, the question appeared after the  
493 last word of the final sentence. Questions were answered using two buttons  
494 that mapped to yes/no ("Ja" / "Nein"), highlighted on the screen in green and  
495 red color, respectively. The position of the correct and incorrect button varied  
496 randomly in order to avoid motor preparation effects.

#### 497 3.1.4 Electrophysiological Recording and Processing

498 The EEG was recorded using 26 active Ag/AgCl electrodes, positioned on the  
499 scalp following the standard 10-20 system. During recording, FCz was used as  
500 online reference and AFz as ground. Data were digitized at a sampling rate of  
501 1000 Hz, leading to a temporal resolution at 1 ms increments. Eye-movement  
502 artefacts were monitored through the electro-oculogram of two electrodes placed  
503 horizontally at the outer canthi of each eye and two electrodes placed vertically  
504 above and below the left eye. We aimed to keep impedances below  $5k\Omega$  on scalp  
505 electrodes and below  $10k\Omega$  on eye electrodes and did not apply online filtering.  
506 We re-referenced the EEG offline to the averages of the left and right mastoid  
507 electrodes and band-pass filtered the data between 0.01 Hz and 30 Hz. Epochs  
508 ranging from -200 to 1200 ms relative to target word onset were extracted from  
509 the EEG signal. Trials with ocular and muscular artefacts were excluded using  
510 a semi-automatic procedure. Baseline correction was performed based on the  
511 200 ms pre-stimulus interval.

#### 512 3.1.5 Analysis

513 To analyse the data, we apply rERPs (Smith & Kutas, 2015), a regression-based  
514 ERP (re-)estimation technique, similar to the analysis used for the self-paced  
515 reading data. For this analysis, we apply linear regression, as opposed to lin-  
516 ear mixed-effects regression, as the analytical solution of solving least-squares  
517 regression will provide stable models and faster computation speed. This will  
518 allow us to re-estimate the data on all electrodes and to inspect topographic dif-  
519 ferences in the analyses. In particular, rERPs apply within-subjects regression  
520 and the models' parameters and forward solutions are averaged across subjects,

521 analogous to the traditional ERP averaging procedure in which condition aver-  
 522 ages are computed from the means of individual subjects. The advantage of the  
 523 rERP technique compared to traditional statistical analyses is that it allows us  
 524 to gauge the relative explanatory power of target word plausibility and distrac-  
 525 tor cloze probability across time and electrodes: The advantage of the rERP  
 526 technique is that by computing a separate regression model for each subject on  
 527 each electrode and time sample, we can trace predictor coefficients, inspect esti-  
 528 mated waveforms and residual error, and obtain effect sizes across the temporal  
 529 and spatial dimensions of the ERP signal. Crucially, this approach goes beyond  
 530 simple condition contrasts, as we are interested in the continuous relationship  
 531 between stimulus properties and ERPs. In fact, the rERP analyses themselves  
 532 are only informed by the continuous by-trial stimulus properties and not by any  
 533 explicit condition coding. That is, we only average by condition after fitting the  
 534 models, to assess the extent to which our predictors capture the effects structure  
 535 across conditions.

536 We will apply the same predictor combination that we used for the analysis  
 537 of the reading times. That is, we model the ERP signal as a function of target  
 538 word plausibility and distractor cloze probability. The model specification for  
 539 the rERP models is

$$Y = \beta_0 + \beta_1 Plaus + \beta_2 Clozedist + \epsilon \quad (2)$$

540 We report coefficients ( $\beta$  terms), estimates (the forward solution  $\hat{Y}$ ), and  
 541 residual error ( $\epsilon$ , the difference between observed data  $Y$  and  $\hat{Y}$ ), averaged  
 542 across subjects. Additionally, we will compute the same model across subjects.  
 543 This has the advantage that we obtain a single t-value and p-value for each  
 544 electrode and time-sample, rather than vectors of t-values and p-values, one for  
 545 each subject. However, there is still a multiple comparisons problem due to  
 546 the multitude of time-samples and electrodes and hence we correct p-values for  
 547 the inflated false discovery rate using the method proposed by Benjamini and  
 548 Hochberg (1995). We adjust p-values separately for the two time windows of  
 549 interest but across nine analysed electrodes and the time-samples within a time  
 550 window (N400: 300 - 500 ms; P600: 600 - 1000 ms).

551    **3.2 Results**

552    **3.2.1 Comprehension Questions**

553 Participants answered comprehension questions on half of the experimental  
554 items. Descriptive metrics for accuracy and reaction times were computed across  
555 subjects. Average Accuracy was 96.2% (SD = 3.9 range = 87.0% - 100.0%).  
556 Mean reaction time was 2162 ms (SD = 254, range = 1568 ms - 2841 ms).  
557 Accuracies and reaction times split per condition are given in Table 8.

Condition	Accuracy			Reaction Time		
	Mean	SD	Range	Mean	SD	Range
A	95.1%	7.3	75.0% - 100.0%	2144 ms	309	1618 ms - 2781 ms
B	98.1%	6.1	75.0% - 100.0%	2153 ms	316	1459 ms - 3077 ms
C	95.5%	8.3	62.5% - 100.0%	2182 ms	325	1522 ms - 2841 ms

Table 8: Task performance on the comprehension questions in the EEG experiment. Accuracy and reaction times were computed across subjects.

558    **3.2.2 ERPs**

559 Grand-averaged ERPs for the three conditions on all non-reference and non-eye  
560 electrodes are displayed in Figure 7. Visual inspection suggests a broadly dis-  
561 tributed negativity, lasting approximately from 250 ms to 400 ms post stimulus  
562 onset in response to target words that are less plausible and for which a seman-  
563 tically attractive alternative interpretation is present (Condition B). A smaller,  
564 more frontally pronounced early negativity, lasting approximately from 250 to  
565 400 ms post stimulus onset, is also evoked by implausible target words (Condi-  
566 tion C) on frontal and central electrodes. Around the typical peak of the N400  
567 component, no pattern of N400 amplitudes with plausibility is observable by vi-  
568 visual inspection. Further, both Condition B (less plausible, semantic attraction)  
569 and Condition C (implausible, no semantic attraction) elicit broadly distributed  
570 positivities, emerging from 500 ms post stimulus onset. The positivity elicited  
571 by Condition C is stronger in amplitude than that elicited by Condition B on  
572 parietal electrodes. On left frontal electrodes, however, their amplitudes appear  
573 similar in parts of the epoch.

574 To further examine the topographies of the condition contrasts, we display  
575 topographic maps of the differences between the conditions in a time window  
576 matching visual inspection of the negativities (250 - 400 ms) and in the canonical

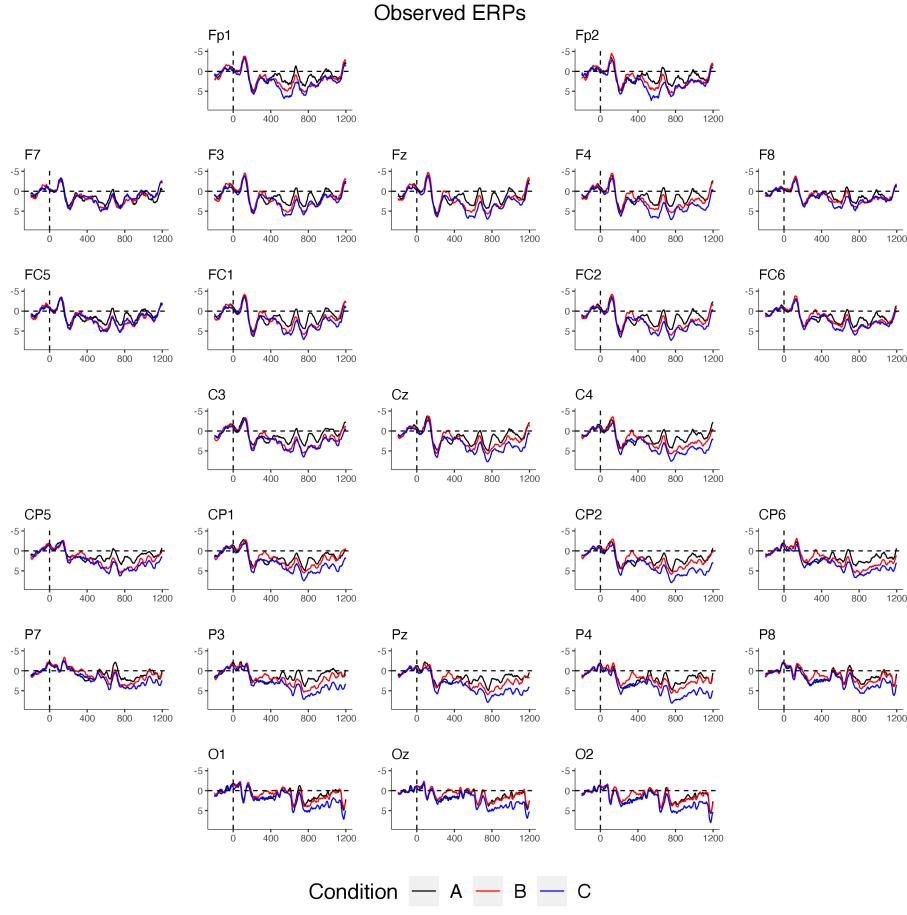


Figure 7: Grand-average ERPs in the three conditions manipulating plausibility and semantic attraction. Negative voltages are plotted upwards.

577 N400 (300 - 400 ms) and P600 time windows (600 - 1000 ms). The topographic  
 578 map of Condition B (less plausible; disconfirmation) relative to Condition A is  
 579 presented in Figure 8. The early negativity is broadly distributed and peaks over  
 580 right parietal electrodes, whereas left-frontally, the difference is smaller. The  
 581 temporal average of the N400 time window exhibits negativities over right pari-  
 582 ental and occipital electrodes. Inspection of the waveforms (Figure 7) strongly  
 583 suggest that this negativity is driven by the temporally overlapping preceding  
 584 negativity and that, additionally, the N400 time window also includes the onset  
 585 of the P600 effect of Condition B relative to A. The late positivity has peaks  
 586 both over left and right central electrodes with a trough between them.

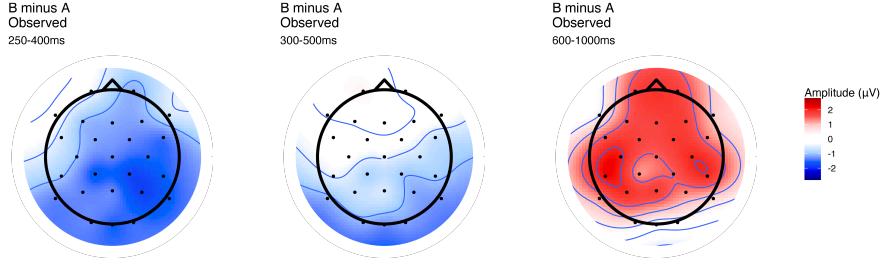


Figure 8: Topographic distributions of the average potentials of Condition B for the earlier negativity (250-400 ms), the canonical N400 (300-500 ms), and P600 (600-1000 ms) time windows, relative to the baseline condition. Topographies are computed from all non-reference and non-eye electrodes.

587 In the topographic maps for Condition C (Figure 9), the early negativity  
 588 appears much smaller than that of Condition B and peaks over left-frontal  
 589 electrodes. The topography in the N400 time window does not contain the  
 590 topography of a typical, centrally peaking N400, but more likely shows the early,  
 591 emerging P600 effect. The late positivity clearly peaks over parietal electrodes.

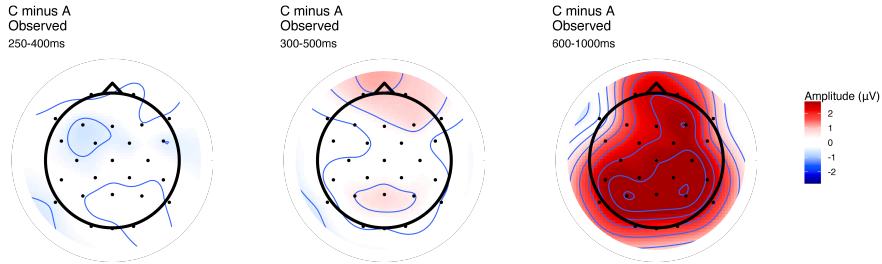


Figure 9: Topographic distributions of the average potentials of Condition C for the earlier negativity (250-400 ms), the canonical N400 (300-500 ms), and P600 (600-1000 ms) time windows, relative to the baseline condition. Topographies are computed from all non-reference and non-eye electrodes.

592 Turning to the rERP analysis, we first inspect the estimated waveforms for  
 593 a single electrode, Pz (Figure 10; left) as well as the residual error (right), i.e.  
 594 the difference between the observed and the estimated data. The estimates  
 595 were generated by a model with target word plausibility and distractor cloze  
 596 probability as predictors. The estimates and residuals suggest that the models  
 597 accurately capture the major trends in the data, as observable by visual inspec-

tion. That is, the models predict a negativity for Condition B between 250 and 400 ms, no negativity for Condition C (on this electrode), and late positivities of increasing amplitudes for Condition B and C, respectively.

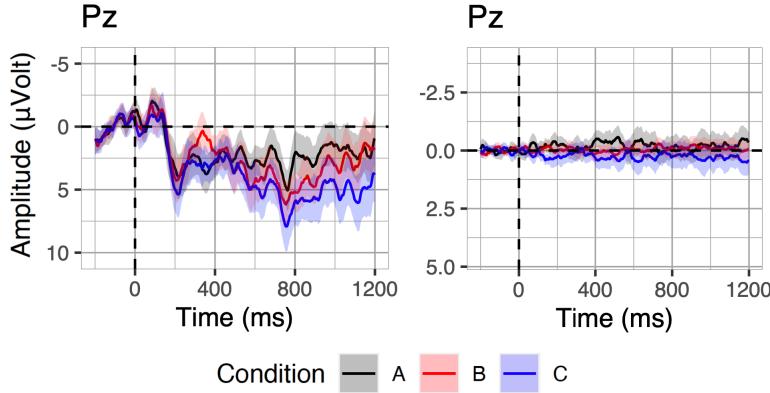


Figure 10: Estimated waveforms (left) and residual error (right) on electrode Pz based on regression models using target word plausibility and distractor cloze probability as predictors.

To assess which predictor captures the voltage deflections, we turn to the model coefficients, plotted over time (Figure 11; right). The coefficient for distractor cloze probability which is high only in Condition B predicts the negativity elicited by that condition. Plausibility, which is stepped across the three conditions, captures the graded late positivities. In order to assess whether distractor cloze probability also predicts a late positivity on another electrode site, we also inspect the coefficients on electrode C3 (Figure 11; left), on which the late positivities for Condition B and C appeared to match (Figure 7). Indeed, on this electrode, distractor cloze probability predicts additional positivity in parts of the P600 time window. On this electrode, plausibility also predicts a smaller earlier negativity.

Using these coefficients, we can now compute the ERPs estimated by a single predictor in isolation. To achieve this, we compute the forward estimates for the entire data set while factoring out the influence of the other predictor by fixing it to its average value, which is zero for z-scored predictors. The isolated estimates of distractor cloze on electrode Pz contain the negativity of Condition B (Figure 12, left). Isolating the estimates of plausibility on electrode Pz (right) reveals no modulation in the N400 time window but the three step modulation in the P600 time window. These estimates suggest that the negativity is elicited

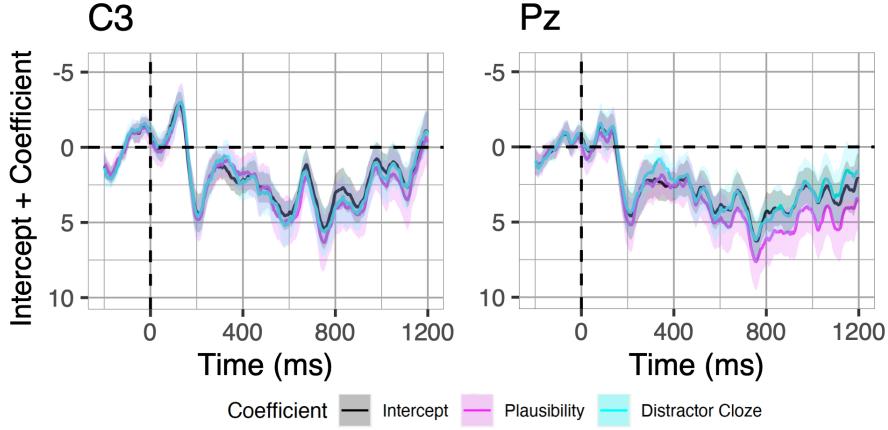


Figure 11: Regression model coefficients (added to their intercept) across time on electrode C3 and Pz.

620 by the expectancy of the distractor word, and that plausibility predicts no N400  
 621 but P600 modulations.

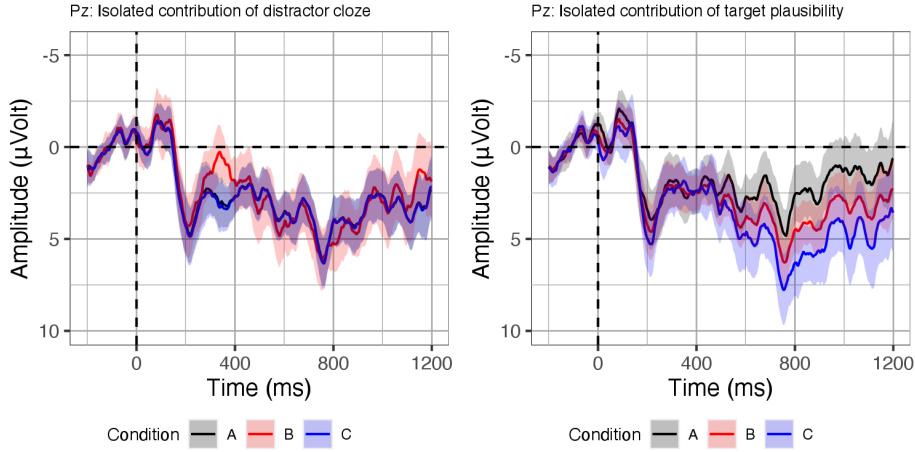


Figure 12: The isolated forward estimates of distractor cloze probability (left) and plausibility (right), based on coefficients that were fitted in models containing both predictors.

622 As the single-electrode inspection of the coefficients suggests potential topo-  
 623 graphic differences between the contributions of the predictors, we visualise  
 624 the estimated ERP data as topographic maps. This allows us to dissect how  
 625 target word plausibility and distractor cloze probability interact in shaping the

626 topographic map of the difference between Condition B and A (see Figure 8).  
 627 Figure 13 displays the individual contributions of distractor cloze probability  
 628 (left), target plausibility (middle left), and how they sum (middle right) to the  
 629 estimated data for Condition B, which is similarly distributed to the observed  
 630 data (right). The topographic maps suggest that while plausibility predicts a  
 631 larger, parietally peaking positivity, there is additional left-frontally peaking  
 632 positivity predicted by distractor cloze probability. This suggests that the over-  
 633 all topographic distribution observed for Condition B (Figure 8) is composed of  
 634 a parietal and a left-frontocentral sub-component.

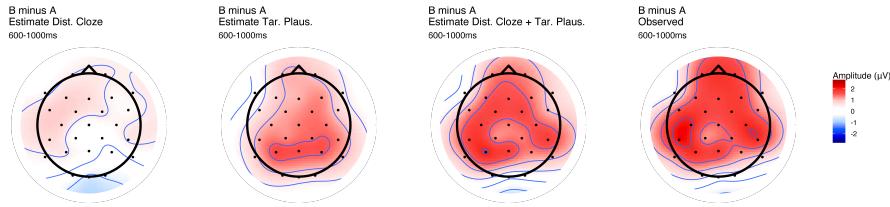


Figure 13: Topographic distributions of the potentials in the P600 time window estimated by distractor cloze probability (left) and plausibility (middle) and their summed estimated potential (right) as well as the observed potential for Condition B between 600-1000, relative to the baseline condition. Topographies computed from all non-reference and non-eye electrodes.

635 To assess the statistical significance of our two predictors, we computed mod-  
 636 els in which we compute the regression coefficients across all subjects, rather  
 637 than fitting individual models per-subject. We report the t-values for the two  
 638 predictors on nine central electrodes (Figure 14). Further, the bar below the t-  
 639 values indicates time-samples that were significant after correcting for multiple  
 640 comparisons within the N400 and the P600 time window and across electrodes  
 641 and time-samples. Our inferential statistics indicate that distractor cloze prob-  
 642 ability significantly predicts a negativity in the 300 ms to 400 ms range. While  
 643 the t-values for plausibility are large on frontal electrodes in the pre-N400 time  
 644 window, indicative of a negativity predicted by low plausibility items, this does  
 645 not reach significance in the current selection of time windows and electrodes.  
 646 Plausibility significantly predicts a late positivity (600 ms - 1000 ms) with a  
 647 peak over parietal electrodes. Distractor cloze probability, while generating a  
 648 left-frontocentral late positivity in the forward estimates (Figure 13), does not  
 649 reach significance in our selection of time windows and electrodes.

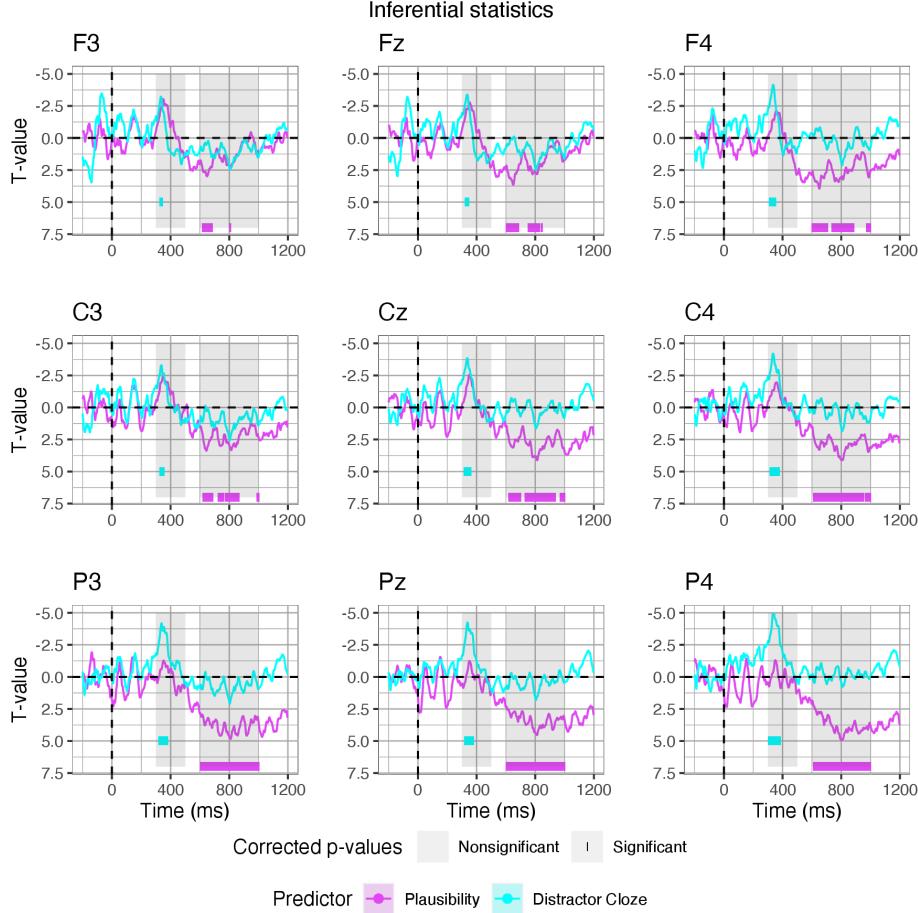


Figure 14: T-values for the predictors plausibility and distractor cloze probability on nine central electrodes based on across-subjects regression. Bars indicate time-samples with significant p-values after multiple comparisons correction.

### 650 3.3 Discussion

651 Experiment 2 replicated the main findings of Nieuwland and van Berkum (2005)  
 652 using visual rather than auditory language comprehension and employing an ex-  
 653 plicit task that incentivises reading for comprehension. In the original design,  
 654 a context paragraph repeatedly mentioned the target words before those same  
 655 words were presented either as plausible or implausible continuations. Rather  
 656 than eliciting an N400, a P600 effect relative to baseline was observed. This  
 657 matches our data in the less plausible condition (B: “Then *weighed* the lady the  
 658 *tourist*”) compared to the baseline (A: “Then *dismissed* the lady the *tourist*”).

659 Further, while a semantically attractive alternative interpretation is globally  
660 available in Condition B, it is unavailable in Condition C (C: “Then *signed*  
661 the lady the tourist”). Indeed, Condition C thus instantiates a classic semantic  
662 incongruity (see Van Petten & Luka, 2012, for a review). On multi-stream  
663 models, the absence of such a semantic attraction (Condition C) should re-  
664 sult in the emergence of an N400 effect compared to the baseline condition.  
665 However, no N400 effect but only a P600 effect was observed in Condition C  
666 relative to A. Further, our design manipulated plausibility on three levels (A:  
667 plausible < B: less plausible < C: implausible), showing that target words with  
668 intermediate plausibility ratings (B: “Then *weighed* the lady the tourist”) also  
669 elicit a P600 effect, intermediate in amplitude, compared to the fully plausible  
670 and implausible conditions. Indeed, the plausibility ratings collected in a  
671 pre-test provided a continuous predictor which significantly predicted the P600  
672 modulations observed across the three conditions.

673 While distractor *absence* did not elicit an N400 effect relative to baseline,  
674 the *presence* of a distractor in fact elicited an earlier negativity, emerging from  
675 around 250 ms and lasting until 400 ms post-stimulus onset for Condition B. An  
676 interpretation of this earlier negativity as an N400 appears implausible given the  
677 temporal invariability of the N400 peak (Federmeier & Laszlo, 2009). Rather,  
678 we interpret this component to be elicited by the strong and unfulfilled expect-  
679 ation of the distractor word on a lexical level. Likely, this early component  
680 often overlaps with the N400 and it is the combination of lexical repetition and  
681 disconfirmation in our experiment that allows us to observe it in isolation. That  
682 is, even though the distractor word was strongly expected and not presented,  
683 lexical retrieval - indexed by the N400 - of the target word’s meaning was still  
684 maximally facilitated. Interestingly, Nieuwland and van Berkum (2005) did not  
685 observe a similar negativity in their study, even though they relied on auditory  
686 presentation - a modality in which a component with a similar time-course, the  
687 phonological mismatch negativity (PMN), is often observed (Connolly, Stewart,  
688 & Phillips, 1990; Hagoort & Brown, 2000; Jachmann, Drenhaus, Staudte, &  
689 Crocker, 2019).

690 Further, our rERP analyses suggest that the presence of a strongly antici-  
691 pated distractor word that is then *not* presented as target word (Condition B)  
692 leads to additional modulation in the late ERP signal with a positive left-frontal  
693 peak. While distractor cloze probability was not significant in the later time  
694 window, a frontal positivity could in fact be expected for our design, as the  
695 way in which our design makes a semantically attractive alternative interpre-

tation available effectively creates a prediction disconfirmation (“Then *weighed* the lady the tourist” where “suitcase” is expected), which has been linked to frontal positivities in previous research (Brothers, Swaab, & Traxler, 2015; DeLong, Quante, & Kutas, 2014; DeLong, Urbach, Groppe, & Kutas, 2011; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Kuperberg, Brothers, & Wlotko, 2020; Quante, Bölte, & Zwitserlood, 2018, see also earlier results by Kutas, 1993). Our rERP analysis suggests that the positivity observed for Condition B can be dissected into two sub-components: A P600 with parietal peak, predicted by plausibility, and a disconfirmation related positivity with left-central positivity peak, predicted by distractor cloze probability. In the design of Nieuwland and van Berkum (2005), a disconfirmation was also present. However, the replacement word was implausible. Their difference waves suggest no apparent deviation from a canonical, parietal P600. This is in line with the finding that the frontal positivity is produced by unexpected but *plausible* target words, whereas unexpected and implausible target words lead to a parietally distributed late positivity (Van Petten & Luka, 2012, for a review see Kuperberg et al., 2020).

## 4 General Discussion

The goal of the present study was to test competing hypotheses about the functional interpretation of N400 and P600 components. In particular, building on a previous study (Nieuwland & van Berkum, 2005) we tested the prediction of RI theory that the P600 is a graded index of integration difficulty (Brouwer et al., 2017; Brouwer, Delogu, Venhuizen, & Crocker, 2021) directly against the predictions made by multi-stream models (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kim & Osterhout, 2005; Kos et al., 2010; Kuperberg, 2007; van Herten et al., 2005, similarly Rabovsky et al., 2018 and Ryskin et al., 2021).

Multi-stream models maintain that the N400 is the index of integration difficulty. On multi-stream accounts, no N400 modulation is generated if the processor initially does not detect an anomaly in the semantic stream because of the availability of a semantically attractive alternative interpretation. The anomaly is then detected by a second, algorithmic stream, and it is the mismatch between the analyses of the semantic stream and the algorithmic stream which produces an increase in P600 amplitude. Under RI theory, by contrast, the N400 is taken to index lexical retrieval. Critically, in our design (see Figure

730 2) — which employs a context manipulation, in which a semantically attractive  
731 alternative is either available or not (Condition B vs. C), and target word plau-  
732 sibility is varied across three levels (Condition A < B < C) — the target word is  
733 repeated several times in a preceding context paragraph. On the retrieval view  
734 of the N400 (Brouwer et al., 2012; Kutas & Federmeier, 2000, 2011; Lau et al.,  
735 2009, 2008; van Berkum, 2009, 2010), this is predicted to maximally facilitate  
736 retrieval of target word meaning and thus minimise N400 differences across con-  
737 ditions. In sum, RI theory predicts no N400 differences across conditions, and  
738 increasing P600 amplitudes as a function of decreasing target word plausibility.  
739 Multi-stream models, on the other hand, predict a P600, but no N400, effect  
740 if a semantically attractive alternative interpretation is available (Condition B  
741 relative to A) and an N400, but no P600, effect if no alternative interpretation  
742 is available (Condition C relative to A).

743 We validated the design in a self-paced reading experiment (Experiment 1)  
744 that revealed a graded sensitivity of reading times to plausibility, indicating  
745 that the stimuli indeed induce graded integration difficulty. Distractor cloze  
746 probability did not modulate reading speed significantly. The EEG experiment  
747 (Experiment 2), replicated the original findings of Nieuwland and van Berkum  
748 (2005), i.e. the absence of an N400 effect and the presence of a P600 effect  
749 for a less plausible relative to a plausible target word when the target word is  
750 strongly primed by the context and in the presence of a semantically attractive  
751 alternative interpretation (our Condition B). Furthermore, our results revealed  
752 the *graded* sensitivity of a posterior late positivity to plausibility, as shown by  
753 stepped P600 amplitudes for plausible (A), less plausible (B), and implausible  
754 (C) items. The absence a plausibility related N400 effect is inconsistent with an  
755 interpretation of the N400 as a graded index of integration difficulty. Addition-  
756 ally, the presence of an expected word which was then not presented elicited an  
757 early negativity (250-400 ms) – likely a correlate of lexical mismatch. Further,  
758 an rERP analysis revealed that the presence of a strongly expected distractor  
759 word – or rather its disconfirmation – resulted in an additional left-frontal pos-  
760 itivity in a later time window, in line with previous research. However, in our  
761 analyses the contribution of disconfirmations to late positivities was not statis-  
762 tically significant. In sum, as we discuss in more detail below, these findings  
763 reveal a critical novel dimension to the functional interpretation of the P600  
764 that has important implications for existing and future neurocognitive experi-  
765 ments and theories, namely that the P600 is a *continuous* index of integration  
766 effort.

767    **4.1 The processing cost of disconfirmed expectations**

768    While the main goal of our design was to manipulate the availability of a se-  
769    mantically attractive alternative interpretation (The lady weighing the suitcase  
770    rather than the tourist), the way in which we achieved this manipulation effec-  
771    tively created a prediction disconfirmation in Condition B. That is, when pre-  
772    senting the final sentence fragment “Then weighed the lady the ...”, “suitcase”  
773    was expected – as shown by high distractor cloze probability – but “tourist” was  
774    encountered instead. While not the main focus of our hypotheses, the results  
775    are relevant to the literature on disconfirmed predictions.

776    For Condition B, we observed an early negativity relative to both Condition  
777    A and C, lasting approximately from 250 to 400 ms post-stimulus onset. This  
778    deflection may relate to the mismatch between the observed word form (target)  
779    and the anticipated word form (distractor). Critically, under the retrieval view  
780    on the N400 (Brouwer et al., 2012; Kutas & Federmeier, 2000, 2011; Lau et al.,  
781    2009, 2008; van Berkum, 2009, 2010), this mismatch does not appear to tax lex-  
782    ical retrieval, as no N400 modulation was observed: The difference between the  
783    waveforms disappeared by 400 ms, which would be the typical peak of the N400  
784    component (Federmeier & Laszlo, 2009). This earlier negative component likely  
785    overlaps with the N400 in previous studies on disconfirmations and it is the ab-  
786    sence of an N400 effect relative to baseline in our data that allows us to observe  
787    the earlier negativity in isolation. Results that are directly relevant to ours  
788    are presented by Brothers et al. (2015), who observed a centrally peaking N250  
789    for the contrast between a medium-cloze unpredicted versus a medium-cloze  
790    predicted target word. Further, in their data, the earlier negativity was not ob-  
791    served for the contrast of a low-cloze unpredicted to a medium-cloze unpredicted  
792    target word, which only elicited an N400 effect. Similarly, the visual mismatch  
793    negativity has been reported for exactly the time window between 250 ms to  
794    400 ms (Tales, Newton, Troscianko, & Butler, 1999). Furthermore, negativities  
795    preceding the N400 time window have been found for expectation-incompatible  
796    relative to expectation-compatible stimuli (Bartholow et al., 2005), for expecta-  
797    tion based semantic priming (Franklin, Dien, Neely, Huber, & Waterson, 2007),  
798    and, using pictorial stimuli, for perceptual hypothesis testing which is argued to  
799    precede multimodal semantic memory access, as indexed by the following N400  
800    (Kumar, Federmeier, & Beck, 2021).

801    In the time window from 600 to 1000 ms, our rERP analysis suggests that  
802    target words that disconfirmed expected distractor words induced a left-frontal

803 positivity. Distractor cloze probability did, however, not reach significance in  
804 the analyses, and hence these results warrant adequate caution. However, pre-  
805 vious research has repeatedly reported frontal positivities elicited by prediction  
806 disconfirmations (Brothers et al., 2015; DeLong et al., 2014, 2011; Federmeier  
807 et al., 2007; Kuperberg et al., 2020; Kutas, 1993; Quante et al., 2018), making  
808 our results relevant to this line of research. A prominent idea has been that  
809 if the target is unexpected but plausible, disconfirmations result in a frontally  
810 pronounced positivity, whereas implausible replacements result in a parietal pos-  
811 itivity (Van Petten & Luka, 2012). We see, however, two open issues with regard  
812 to this strict functional segregation of frontal and parietal positivities. First,  
813 the apparent distinction between frontally and parietally distributed positivi-  
814 ties could be an artefact of spatio-temporal component overlap with the N400  
815 (Brouwer & Crocker, 2017; Delogu, Brouwer, & Crocker, 2021), and secondly,  
816 frontally and parietally distributed positivties may not be mutually exclusive.

817 A relevant study by DeLong et al. (2014) included plausible, less plausible  
818 disconfirming, and implausible disconfirming target words. The design elicited  
819 a frontal positivity for less plausible disconfirming words, a parietal positivity  
820 for implausible disconfirming words and, critically, N400 effects in response to  
821 both less plausible and implausible words, relative to baseline. Our design, on  
822 the other hand, does not elicit N400 differences and hence circumvents the issue  
823 of component overlap, thereby providing a clearer view on the distribution of  
824 the late positivities. The estimates generated by our rERP models (Figure 13)  
825 suggest that even without a strong N400 overlapping with the late positivity,  
826 unfulfilled expectations create an additional positivity with a left-frontocentral  
827 distribution. Further, in the disconfirming condition (B), the context addition-  
828 ally made the target word less plausible compared to the baseline condition.  
829 Our rERP analysis revealed that for Condition B, plausibility induces a pari-  
830 etal P600 – which was not observed in the data of DeLong et al. (2014) – in  
831 addition to the frontal positivity elicited by the disconfirmation. In sum, our  
832 results and the rERP analysis suggests that disconfirmations indeed induce a  
833 frontal positivity, but that this frontal positivity can co-occur with a plausibil-  
834 ity related parietal positivity on less plausible, but ultimately possible target  
835 words.

836    **4.2 Global revision on the multi-stream account**

837    The main goal of this study was to test the hypotheses of multi-stream models  
838    against those of RI theory. Multi-stream models were originally proposed in  
839    response to studies eliciting “semantic P600s”, in which semantic anomalies did  
840    not elicit N400 but rather P600 effects relative to baseline. Multi-stream ac-  
841    counts explain some of the original data points, by postulating that the semantic  
842    stream does not detect the anomaly because a semantically attractive alterna-  
843    tive interpretation is available. For instance, in order to “repair” the sentence  
844    “the hearty meal was devouring”, the inflection of the verb could be changed to  
845    “devoured”, yielding a plausible interpretation. However, the surface structure  
846    of the sentence does not match this interpretation, which is detected by the  
847    algorithmic stream and the conflict between the two streams leads to a P600  
848    effect when compared to a congruous condition.

849    This explanation was based on a *locally* available alternative interpretation  
850    (see Figure 1). However, no such *local* availability is given in the design of  
851    Nieuwland and van Berkum (2005, “Next, the lady told the tourist/suitcase”),  
852    and accordingly, an N400 and no P600 effect relative to baseline would be pre-  
853    dicted by multi-stream models. However, the reverse pattern was observed.  
854    To account for this, multi-stream may invoke a *globally* attractive alternative  
855    interpretation. That is, making use of the *globally* available information, the  
856    word “suitcase” could be replaced with the discourse-salient word “tourist” in  
857    order to arrive at a plausible interpretation in the semantic stream. Again,  
858    the analysis generated by the algorithmic stream conflicts with the analysis of  
859    the semantic stream, explaining the P600 found by Nieuwland and van Berkum  
860    (2005). Importantly, it follows that if neither a *locally* nor a *globally* available  
861    alternative interpretation is present, an N400 effect should be observed relative  
862    to baseline.

863    The current study extended the original design by Nieuwland and van Berkum  
864    (2005) to test this prediction. In the new context manipulation design, we made  
865    an alternative interpretation available *globally* for a less plausible target word  
866    (Condition B: “Next, the lady weighed the tourist”), whereas no alternative  
867    interpretation was available for the fully implausible target word (Condition C:  
868    “Next, the lady signed the tourist”). Assuming the *globally* active plausibility  
869    heuristic described above, multi-stream models predict only a P600 effect for  
870    Condition B and only an N400 effect for Condition C relative to Condition A.  
871    Note that multi-stream models in general predict either an N400 or a P600,

872 which makes all bi-phasic N400-P600 results problematic for multi-stream ac-  
873 counts (see Van Petten & Luka, 2012, for an overview and Brouwer et al., 2012  
874 for discussion).

875 In Condition B, for which only a P600 is predicted by multi-stream accounts,  
876 we found a P600 effect relative to Condition A. This condition replicates the  
877 results of Nieuwland and van Berkum (2005), and, accordingly, multi-stream  
878 models can only explain this P600 effect by invoking a *globally* available alter-  
879 native interpretation. In Condition C, for which only an N400 effect is predicted  
880 by multi-stream accounts, we observed only a P600 effect relative to Condition  
881 A. Critically, the absence of an N400 effect relative to baseline when any seman-  
882 tically attractive alternative interpretation is removed provides strong evidence  
883 against multi-stream accounts. One explanation of the absence of the N400 ef-  
884 ffect in Condition C relative to A would be to assume that the revision process  
885 changed the context of Condition C (“Then *signed* the lady the”) to make the  
886 target word (“tourist”) plausible. It is difficult, however, to imagine a mecha-  
887 nism that could revise the context in such a way, while at the same time predict-  
888 ing the presence of N400 effects in cases of canonical semantic incongruencies  
889 (see Van Petten & Luka, 2012). Another explanation would entail misunder-  
890 standing “tourist” for something contextually relevant, such as the “tourist’s  
891 ticket”. Many of our stimuli, however, contain strong selectional restriction vio-  
892 lations, such as “the apprentice ate the hammer” (see Appendix A), where such  
893 an explanation does not apply, and hence this explanation cannot account for  
894 the complete absence of an N400 effect of Condition C relative to A. Again, it  
895 is difficult to see how an account would predict absence of N400 effect for the  
896 present stimuli, while at the same predicting the presence of an N400 effect for  
897 canonical semantic incongruencies. In sum, we do not see how the present data  
898 can be reconciled with the mechanisms assumed by multi-stream accounts.

### 899 4.3 Retrieval facilitation under repetition priming

900 The current design had the goal of maximally priming the target word by men-  
901 tioning it repeatedly in a context paragraph preceding the final sentence. The  
902 prediction of RI theory was that maximal priming should maximally facilitate  
903 retrieval of the target word’s meaning from long term memory, thus leading  
904 to equal N400 amplitudes across conditions. Our results revealed that while  
905 an earlier negativity was present in Condition B relative to A (see above), no  
906 difference in the canonical N400 time window was observed for any condition

907 contrast - in line with the retrieval view of the N400 (Brouwer et al., 2012;  
908 Kutas & Federmeier, 2000, 2011; Lau et al., 2009, 2008; van Berkum, 2009,  
909 2010). This study thus adds to several studies that elicited no N400 differences  
910 for target words that were equally strongly or weakly primed by the preceding  
911 context (Delogu et al., 2019, 2021; Hoeks et al., 2004; Kim & Osterhout, 2005;  
912 Kos et al., 2010; Kuperberg, 2007; Nieuwland & van Berkum, 2005; Otten &  
913 van Berkum, 2008; van Herten et al., 2005).

914 Critically, our results show that even when the target word is of intermediate  
915 plausibility (Condition B) or entirely implausible (Condition C), no N400 is pro-  
916 duced – a result that is at odds with the traditional interpretation of the N400  
917 as semantic integration (Brown & Hagoort, 1993, 2000; Hagoort et al., 2004).  
918 Furthermore, it is also inconsistent with hybrid views that assume the N400 to  
919 index both retrieval and integration (Baggio & Hagoort, 2011; Nieuwland et al.,  
920 2020). That is, even though retrieval may be facilitated, these accounts still  
921 predict increased integration effort to be reflected in the N400. Nieuwland et  
922 al. (2020), for instance, suggest that the earlier part of the N400 is sensitive  
923 to retrieval processes, while the later part indexes integration. Critically, how-  
924 ever, we did not observe any N400 differences in either the earlier or later part  
925 of this component, thereby also ruling out hybrid views on the N400. On a  
926 final note, the absence of N400 modulations by plausibility supports the view  
927 that the correlation between corpus-based word surprisal and the N400 may be  
928 best explained by expectation-based modulations of lexical retrieval rather than  
929 integration (see Aurnhammer et al., 2021; Frank et al., 2015, for discussion).

#### 930 4.4 The P600 as a graded index of integration difficulty

931 Most strikingly, our ERP data revealed an important novel dimension of the  
932 P600 component: Our design manipulated plausibility on three levels (plau-  
933 sible, less plausible, implausible) and revealed that P600 amplitude patterns  
934 with plausibility. Going beyond the three discrete levels of plausibility, we suc-  
935 cessfully modeled the ERP signal as a continuous function of numeric per-item  
936 plausibility ratings collected in a pre-test, indicating that the P600 may indeed  
937 be a continuous index of integration difficulty. We conclude that P600s are  
938 not only elicited by highly implausible, impossible, or violating target words  
939 (Bornkessel-Schlesewsky et al., 2011; Kuperberg, 2007), but rather, that P600  
940 amplitude is modulated as a function of integration difficulty by every word.

941 Our proposition that the P600 is a continuous index of integration difficulty

942 is indeed supported by numerous previous studies showing P600 effects for non-  
943 violating but semantically or pragmatically taxing continuations (Burkhardt,  
944 2006, 2007; Cohn & Kutas, 2015; Delogu et al., 2019; Dimitrova et al., 2012;  
945 Hoeks et al., 2013; Regel et al., 2010; Schumacher, 2011; Spotorno et al., 2013;  
946 Xu & Zhou, 2016). For instance, a world knowledge implausibility without a  
947 violation of selectional restrictions, induced a P600 effect relative to control  
948 (Delogu et al., 2019). The graded nature of the P600 was also suggested by a  
949 post-hoc analysis conducted by Aurnhammer et al. (2021). By analysing the  
950 data of the baseline condition only (“Yesterday sharpened the lumberjack [...]  
951 the axe”, translated from German), it was found that not only the N400 but  
952 also the P600 varied gradually as a function of target word expectancy. This  
953 data point was interpreted as indicating a gradual modulation of lexical retrieval  
954 (N400) and integration (P600) as a function of word expectancy. Hence, the  
955 current study directly supports their exploratory, post-hoc analysis with regard  
956 to the P600 component.

957 In Experiment 1, the observed reading times closely patterned with the  
958 P600s in that both were modulated by plausibility on the three levels of our  
959 manipulation. Taken together with the absence of N400 modulations by plau-  
960 sibility, this strengthens the proposed link between reading times and the P600  
961 through Comprehension-centric Surprisal (Brouwer, Delogu, Venhuizen, & Crocker,  
962 2021). To further test this idea, we conduct a post-hoc analysis, in which we  
963 apply the rERP technique to model the ERPs obtained in Experiment 2 by  
964 the per-item reading times obtained in Experiment 1. The resulting coefficients  
965 (Figure 15) suggest that indeed, the observed positivities are correlated to the  
966 observed reading times, suggesting they may be closely associated indices of pro-  
967 cessing difficulty across pools of participants. This finding further corroborates  
968 the P600 as a continuous index of integration effort.

## 969 5 Conclusion

970 Event-Related Potentials provide a multi-dimensional window into the nature  
971 and time-course of language comprehension. Critically, establishing the locus  
972 of specific sub-processes of comprehension in the ERP signal has direct conse-  
973 quences for our understanding of the temporal organisation and architecture of  
974 the comprehension system. The present study directly tested competing views  
975 on whether the N400 or the P600 component of the ERP signal indexes the in-

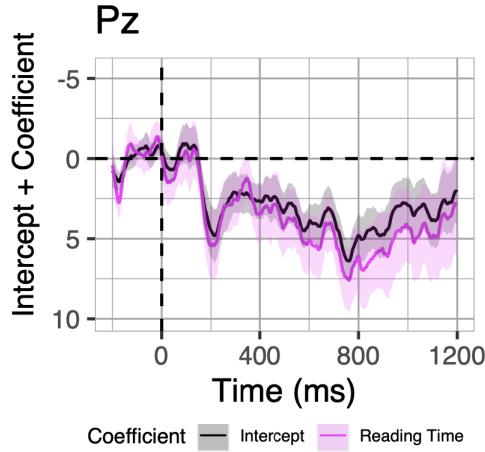


Figure 15: Regression model coefficients (added to their intercept) across time on electrode Pz from models predicting the ERP as a function of the per-item reading times obtained in Experiment 1.

976 tegration of incoming word meaning into an unfolding utterance representation.  
 977 Crucially, the traditional view of the N400 as an index of integration relies on  
 978 the presence of a semantically attractive alternative interpretation to explain  
 979 the absence of an N400 effect in response to certain semantic anomalies. The  
 980 more recent view of the P600 as an index of integration, in turn, predicts P600  
 981 amplitude to be a continuous index of integration difficulty, a prediction that  
 982 had yet to be confirmed. We harnessed these predictions to decide between the  
 983 competing views using a design in which a semantically attractive alternative is  
 984 either available or not, and target word plausibility is varied across three lev-  
 985 els. Further, to minimise lexical processing differences across conditions, target  
 986 words were equally primed by the prior context.

987 An initial self-paced reading study revealed a gradual slow-down of reading  
 988 times for gradual decreases in target word plausibility, suggesting differential  
 989 integration difficulty. In the ERP study, the plausibility manipulation did not  
 990 elicit any N400 differences across conditions. Indeed, the lack of an increased  
 991 N400 for the implausible conditions – even when no semantically attractive  
 992 alternative interpretation is available – is directly at odds with the prediction  
 993 made by contemporary models that maintain the N400 as an index of semantics-  
 994 driven, “quasi-compositional” integration. In fact, the plausibility manipulation  
 995 rather revealed P600 amplitude to be graded for plausibility. Taken together,

these results cannot be reconciled with the N400 as an index of integration, while they are consistent with the P600 as a continuous index of integrative effort. More generally, the results are consistent with Retrieval-Integration theory, a single-stream account in which the N400 indexes lexical retrieval from long-term memory and the P600 indexes integration of incoming word meaning into an unfolding utterance representation. No N400 differences were found, as lexical retrieval was equally facilitated across conditions through repetition priming, and the link between plausibility, reading times, and P600 amplitude establishes the P600 as a direct index of semantic integration that – in line with a comprehension-centric notion of surprisal – is continuous in amplitude as a function of integration effort. This novel dimension of the P600 has important implications for existing and future experiments, as well as for theories and models of language comprehension.

## 6 Acknowledgements

We would like to thank Mante Nieuwland for providing the materials of the Nieuwland and van Berkum (2005) design. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 232722074—SFB 1102.

## References

- 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), e0257430. doi: <https://doi.org/10.1371/journal.pone.0257430>
- 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. doi: <https://doi.org/10.1080/01690965.2010.542671>
- Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and medial frontal negativity: Beyond errors and response conflict. *Psychophysiology*, 42(1), 33–42. doi: <https://doi.org/10.1111/j.1469-8986.2005.00258.x>

- 1027 Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate:  
1028     A practical and powerful approach to multiple testing. *Journal of the  
1029     Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. doi:  
1030     <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- 1031 Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç,  
1032     S., Philipp, M., ... Schlesewsky, M. (2011). Think globally:  
1033     Cross-linguistic variation in electrophysiological activity during sen-  
1034     tence comprehension. *Brain and Language*, 117(3), 133–152. doi:  
1035     <https://doi.org/10.1016/j.bandl.2010.09.010>
- 1036 Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An al-  
1037     ternative perspective on “semantic P600” effects in language  
1038     comprehension. *Brain Research Reviews*, 59(1), 55–73. doi:  
1039     <https://doi.org/10.1016/j.brainresrev.2008.05.003>
- 1040 Brothers, T., Swaab, T. Y., & Traxler, M. J. (2015). Ef-  
1041     fects of prediction and contextual support on lexical processing:  
1042     Prediction takes precedence. *Cognition*, 136, 135–149. doi:  
1043     <https://doi.org/10.1016/j.cognition.2014.10.017>
- 1044 Brouwer, H., & Crocker, M. W. (2017). On the proper treatment of the N400  
1045     and P600 in language comprehension. *Frontiers in Psychology*, 8, 1327.  
1046     doi: <https://doi.org/10.3389/fpsyg.2017.01327>
- 1047 Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. J.  
1048     (2017). A neurocomputational model of the N400 and the P600  
1049     in language processing. *Cognitive Science*, 41, 1318–1352. doi:  
1050     <https://doi.org/10.1111/cogs.12461>
- 1051 Brouwer, H., Delogu, F., & Crocker, M. W. (2021). Splitting event-related  
1052     potentials: Modeling latent components using regression-based wave-  
1053     form estimation. *European Journal of Neuroscience*, 53, 974–995. doi:  
1054     <https://doi.org/10.1111/ejn.14961>
- 1055 Brouwer, H., Delogu, F., Venhuizen, N. J., & Crocker, M. W. (2021). Neu-  
1056     robehavioral correlates of surprisal in language comprehension: A neu-  
1057     rocomputational model. *Frontiers in Psychology*, 12, 615538. doi:  
1058     <https://doi.org/10.3389/fpsyg.2021.615538>
- 1059 Brouwer, H., Fitz, H., & Hoeks, J. C. J. (2012). Getting real about  
1060     Semantic Illusions: Rethinking the functional role of the P600 in  
1061     language comprehension. *Brain Research*, 1446, 127–143. doi:  
1062     <https://doi.org/10.1016/j.brainres.2012.01.055>
- 1063 Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence

- 1064 from masked priming. *Journal of Cognitive Neuroscience*, 5(1), 34–44.  
1065 doi: <https://doi.org/10.1162/jocn.1993.5.1.34>
- 1066 Brown, C., & Hagoort, P. (2000). On the electrophysiology of language comprehen-  
1067 sion: Implications for the human language system. In M. W. Crocker,  
1068 P. M., & C. C. Jr. (Eds.), *Architectures and mechanisms for language*  
1069 *processing* (pp. 213–237). Cambridge University Press.
- 1070 Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mecha-  
1071 nisms: Evidence from event-related brain potentials. *Brain and Language*,  
1072 98(2), 159–168. doi: <https://doi.org/10.1016/j.bandl.2006.04.005>
- 1073 Burkhardt, P. (2007). The P600 reflects cost of new informa-  
1074 tion in discourse memory. *NeuroReport*, 18(17), 1851–1854. doi:  
1075 <https://doi.org/10.1097/WNR.0b013e3282f1a999>
- 1076 Cohn, N., & Kutas, M. (2015). Getting a cue before get-  
1077 ting a clue: Event-related potentials to inference in visual nar-  
1078 rative comprehension. *Neuropsychologia*, 77, 267–278. doi:  
1079 <https://doi.org/10.1016/j.neuropsychologia.2015.08.026>
- 1080 Connolly, J. F., Stewart, S. H., & Phillips, N. A. (1990). The effects of pro-  
1081 cessing requirements on neurophysiological responses to spoken sentences.  
1082 *Brain and Language*, 39(2), 302–318. doi: [https://doi.org/10.1016/0093-934X\(90\)90016-A](https://doi.org/10.1016/0093-934X(90)90016-A)
- 1083 Delogu, F., Brouwer, H., & Crocker, M. W. (2019). Event-related po-  
1084 tentials index lexical retrieval (N400) and integration (P600) during  
1085 language comprehension. *Brain and Cognition*, 135, 103569. doi:  
1086 <https://doi.org/10.1016/j.bandc.2019.05.007>
- 1087 Delogu, F., Brouwer, H., & Crocker, M. W. (2021). When com-  
1088 ponents collide: Spatiotemporal overlap of the N400 and P600 in  
1089 language comprehension. *Brain Research*, 1766, 147514. doi:  
1090 <https://doi.org/10.1016/j.brainres.2021.147514>
- 1091 DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictabil-  
1092 ity, plausibility, and two late ERP positivities during written  
1093 sentence comprehension. *Neuropsychologia*, 61, 150–162. doi:  
1094 <https://doi.org/10.1016/j.neuropsychologia.2014.06.016>
- 1095 DeLong, K. A., Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlap-  
1096 ping dual ERP responses to low cloze probability sentence continuations.  
1097 *Psychophysiology*, 48(9), 1203–1207. doi: <https://doi.org/10.1111/j.1469-8986.2011.01199.x>
- 1098 Dimitrova, D. V., Stowe, L. A., Redeker, G., & Hoeks, J. C. J. (2012). Less  
1099
- 1100

- 1101 Is Not More: Neural Responses to Missing and Superfluous Accents in  
1102 Context. *Journal of Cognitive Neuroscience*, 24(12), 2400–2418. doi:  
1103 [https://doi.org/10.1162/jocn\\_a.00302](https://doi.org/10.1162/jocn_a.00302)
- 1104 Eddine, S. N., Brothers, T., & Kuperberg, G. R. (2022). The N400 in silico:  
1105 A review of computational models. *in preparation*. (Publisher: Academic  
1106 Press)
- 1107 Federmeier, K. D., & Laszlo, S. (2009). Chapter 1 Time for Meaning: Elec-  
1108 trophysiology Provides Insights into the Dynamics of Representation and  
1109 Processing in Semantic Memory. In *Psychology of Learning and Motivation*  
1110 (Vol. 51, pp. 1–44). Academic Press. doi: [https://doi.org/10.1016/S0079-7421\(09\)51001-8](https://doi.org/10.1016/S0079-7421(09)51001-8)
- 1111 Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007).  
1112 Multiple effects of sentential constraint on word processing. *Brain Re-  
1113 search*, 1146, 75–84. doi: <https://doi.org/10.1016/j.brainres.2006.06.101>
- 1114 Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response  
1115 to the amount of information conveyed by words in sentences. *Brain and  
1116 Language*, 140, 1–11. doi: <https://doi.org/10.1016/j.bandl.2014.10.006>
- 1117 Franklin, M. S., Dien, J., Neely, J. H., Huber, E., & Waterson,  
1118 L. D. (2007). Semantic priming modulates the N400, N300,  
1119 and N400RP. *Clinical Neurophysiology*, 118(5), 1053–1068. doi:  
1120 <https://doi.org/10.1016/j.clinph.2007.01.012>
- 1121 Frisson, S., Harvey, D. R., & Staub, A. (2017). No prediction error cost in read-  
1122 ing: Evidence from eye movements. *Journal of Memory and Language*,  
1123 95, 200–214. doi: <https://doi.org/10.1016/j.jml.2017.04.007>
- 1124 Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntac-  
1125 tic positive shift (SPS) as an ERP measure of syntactic process-  
1126 ing. *Language and Cognitive Processes*, 8(4), 439–483. doi:  
1127 <https://doi.org/10.1080/01690969308407585>
- 1128 Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech:  
1129 semantic ERP effects. *Neuropsychologia*, 38(11), 1518–1530. doi:  
1130 [https://doi.org/10.1016/S0028-3932\(00\)00052-X](https://doi.org/10.1016/S0028-3932(00)00052-X)
- 1131 Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). In-  
1132 tegration of word meaning and world knowledge in language compre-  
1133 hension. *Science*, 304(5669), 438–441. (Publisher: American As-  
1134 sociation for the Advancement of Science Section: Report) doi:  
1135 <https://doi.org/10.1126/science.1095455>
- 1136 Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context:

- 1138        The interaction of lexical and sentence level information during reading.  
1139        *Cognitive Brain Research*, 19(1), 59–73.
- 1140        Hoeks, J. C. J., Stowe, L. A., Hendriks, P., & Brouwer, H. (2013). Questions  
1141        left unanswered: How the brain responds to missing information. *PLOS  
1142        ONE*, 8(10), e73594. doi: <https://doi.org/10.1371/journal.pone.0073594>
- 1143        Jachmann, T. K., Drenhaus, H., Staudte, M., & Crocker, M. W. (2019). In-  
1144        fluence of speakers' gaze on situated language comprehension: Evidence  
1145        from Event-Related Potentials. *Brain and Cognition*, 135, 103571. doi:  
1146        <https://doi.org/10.1016/j.bandc.2019.05.009>
- 1147        Kim, A., & Osterhout, L. (2005). The independence of com-  
1148        binatory semantic processing: Evidence from event-related poten-  
1149        tials. *Journal of Memory and Language*, 52(2), 205–225. doi:  
1150        <https://doi.org/10.1016/j.jml.2004.10.002>
- 1151        Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003).  
1152        Structure and limited capacity in verbal working memory: A study  
1153        with event-related potentials. *Brain and Language*, 85(1), 1–36. doi:  
1154        [https://doi.org/10.1016/S0093-934X\(02\)00548-5](https://doi.org/10.1016/S0093-934X(02)00548-5)
- 1155        Kos, M., Vosse, T. G., Van Den Brink, D., & Hagoort, P. (2010). About  
1156        Edible Restaurants: Conflicts between Syntax and Semantics as Revealed  
1157        by ERPs. *Frontiers in Psychology*, 1.
- 1158        Kumar, M., Federmeier, K. D., & Beck, D. M. (2021). The N300:  
1159        An Index for Predictive Coding of Complex Visual Objects and  
1160        Scenes. *Cerebral Cortex Communications*, 2(2), ttab030. doi:  
1161        <https://doi.org/10.1093/texcom/ttab030>
- 1162        Kuperberg, G. R. (2007). Neural mechanisms of language comprehen-  
1163        sion: Challenges to syntax. *Brain Research*, 1146, 23–49. doi:  
1164        <https://doi.org/10.1016/j.brainres.2006.12.063>
- 1165        Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of  
1166        two positivities and the N400: Distinct neural signatures are evoked  
1167        by confirmed and violated predictions at different levels of repre-  
1168        sentation. *Journal of Cognitive Neuroscience*, 32(1), 12–35. doi:  
1169        [https://doi.org/10.1162/jocn\\_a.01465](https://doi.org/10.1162/jocn_a.01465)
- 1170        Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale,  
1171        A. M., & Caplan, D. (2003). Distinct Patterns of Neural Mod-  
1172        ulation during the Processing of Conceptual and Syntactic Anoma-  
1173        lies. *Journal of Cognitive Neuroscience*, 15(2), 272–293. doi:  
1174        <https://doi.org/10.1162/089892903321208204>

- 1175 Kutas, M. (1993). In the company of other words: Electrophysiological evidence  
1176 for single-word and sentence context effects. *Language and Cognitive Processes*, 8(4), 533–572. doi: <https://doi.org/10.1080/01690969308407587>
- 1177
- 1178 Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic  
1179 memory use in language comprehension. *Trends in Cognitive Sciences*,  
1180 4(12), 463–470. doi: [https://doi.org/10.1016/S1364-6613\(00\)01560-6](https://doi.org/10.1016/S1364-6613(00)01560-6)
- 1181 Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Find-  
1182 ing meaning in the N400 component of the event-related brain po-  
1183 tential (ERP). *Annual Review of Psychology*, 62, 621–647. doi:  
1184 <https://doi.org/10.1146/annurev.psych.093008.131123>
- 1185 Lau, E., Almeida, D., Hines, P. C., & Poeppel, D. (2009). A lexical basis for  
1186 N400 context effects: Evidence from MEG. *Brain and Language*, 111(3),  
1187 161–172. doi: <https://doi.org/10.1016/j.bandl.2009.08.007>
- 1188 Lau, E., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:  
1189 (De)constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933.  
1190 doi: <https://doi.org/10.1038/nrn2532>
- 1191 Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*,  
1192 106(3), 1126–1177. doi: <https://doi.org/10.1016/j.cognition.2007.05.006>
- 1193 Luke, S. G., & Christianson, K. (2016). Limits on lexical pre-  
1194 diction during reading. *Cognitive Psychology*, 88, 22–60. doi:  
1195 <https://doi.org/10.1016/j.cogpsych.2016.06.002>
- 1196 Monsalve, I. F., Frank, S. L., & Vigliocco, G. (2012). Lexical Surprisal As a  
1197 General Predictor of Reading Time. In *Proceedings of the 13th Conference*  
1198 *of the European Chapter of the Association for Computational Linguistics*  
1199 (pp. 398–408). Stroudsburg, PA, USA: Association for Computational  
1200 Linguistics.
- 1201 Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E.,  
1202 Donaldson, D. I., ... Von Grebmer Zu Wolfsturn, S. (2020). Dissociable  
1203 effects of prediction and integration during language comprehension: Evi-  
1204 dence from a large-scale study using brain potentials. *Philosophical Trans-  
1205 actions of the Royal Society B: Biological Sciences*, 375(1791), 20180522.  
1206 doi: <https://doi.org/10.1098/rstb.2018.0522>
- 1207 Nieuwland, M. S., & van Berkum, J. J. A. (2005). Testing the limits of the  
1208 semantic illusion phenomenon: ERPs reveal temporary semantic change  
1209 deafness in discourse comprehension. *Cognitive Brain Research*, 24(3),  
1210 691–701. doi: <https://doi.org/10.1016/j.cogbrainres.2005.04.003>
- 1211 Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited

- 1212 by syntactic anomaly. *Journal of Memory and Language*, 31(6), 785–806.  
1213 doi: [https://doi.org/10.1016/0749-596X\(92\)90039-Z](https://doi.org/10.1016/0749-596X(92)90039-Z)
- 1214 Otten, M., & van Berkum, J. J. A. (2008). Discourse-based word anticipation  
1215 during language processing: Prediction or priming? *Discourse Processes*,  
1216 45(6), 464–496. doi: <https://doi.org/10.1080/01638530802356463>
- 1217 Quante, L., Bölte, J., & Zwitserlood, P. (2018). Dissociating predictability,  
1218 plausibility and possibility of sentence continuations in reading: evidence  
1219 from late-positivity ERPs. *PeerJ*, 6, e5717. (Publisher: PeerJ Inc.) doi:  
1220 <https://doi.org/10.7717/peerj.5717>
- 1221 Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Mod-  
1222 elling the N400 brain potential as change in a probabilistic represen-  
1223 tation of meaning. *Nature Human Behaviour*, 2(9), 693–705. doi:  
1224 <https://doi.org/10.1038/s41562-018-0406-4>
- 1225 Rabovsky, M., & McClelland, J. L. (2020). Quasi-compositional mapping from  
1226 form to meaning: a neural network-based approach to capturing neural  
1227 responses during human language comprehension. *Philosophical Transac-  
1228 tions of the Royal Society B: Biological Sciences*, 375(1791), 20190313.  
1229 doi: 10.1098/rstb.2019.0313
- 1230 Regel, S., Gunter, T. C., & Friederici, A. D. (2010). Isn't it  
1231 ironic? An electrophysiological exploration of figurative language pro-  
1232 cessing. *Journal of Cognitive Neuroscience*, 23(2), 277–293. doi:  
1233 <https://doi.org/10.1162/jocn.2010.21411>
- 1234 Rich, S., & Harris, J. (2021). Unexpected guests: When disconfirmed predictions  
1235 linger. In *Proceedings of the annual meeting of the cognitive science society*  
1236 (pp. 2246–2252). Cognitive Science Society.
- 1237 Ryskin, R., Stearns, L., Bergen, L., Eddy, M., Fedorenko, E., & Gibson, E.  
1238 (2021). An ERP index of real-time error correction within a noisy-channel  
1239 framework of human communication. *Neuropsychologia*, 158, 107855. doi:  
1240 <https://doi.org/10.1016/j.neuropsychologia.2021.107855>
- 1241 Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime: User's guide*.  
1242 Psychology Software Incorporated.
- 1243 Schumacher, P. B. (2011). The hepatitis called...: Electrophysiological evidence  
1244 for enriched composition. In M. J & S. M (Eds.), *Experimental pragmat-  
1245 ics/semantics* (p. 199-2019). John Benjamins Publishing, Amsterdam,  
1246 NL.
- 1247 Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the  
1248 scope of facilitation for upcoming words in sentences. *Journal of Exper-*

- 1249 *imental Psychology: Learning, Memory, and Cognition*, 14(2), 344–354.  
1250 doi: <https://doi.org/10.1037/0278-7393.14.2.344>

1251 Smith, N. J., & Kutas, M. (2015). Regression-based estimation of ERP wave-  
1252 forms: I. The rERP framework. *Psychophysiology*, 52(2), 157–168. doi:  
1253 <https://doi.org/10.1111/psyp.12317>

1254 Spotorno, N., Cheylus, A., Henst, J.-B. V. D., & Noveck, I. A. (2013). What's  
1255 behind a P600? Integration operations during irony processing. *PLOS  
1256 ONE*, 8(6), e66839. doi: <https://doi.org/10.1371/journal.pone.0066839>

1257 Tales, A., Newton, P., Troscianko, T., & Butler, S. (1999). Mismatch neg-  
1258 ativity in the visual modality. *Neuroreport*, 10(16), 3363–3367. doi:  
1259 <https://doi.org/10.1097/00001756-199911080-00020>

1260 Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., &  
1261 Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental  
1262 Psychology: Learning, Memory, and Cognition*, 31(3), 443. doi:  
1263 <https://psycnet.apa.org/doi/10.1037/0278-7393.31.3.443>

1264 van Berkum, J. J. A. (2009). The neuropragmatics of 'simple' utterance compre-  
1265 hension: An ERP review. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and pragmatics: From experiment to theory* (pp. 276–316). Palgrave  
1266 Macmillan, New York.

1267 van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about  
1268 good and bad - Any implications for neuropragmatics? *Italian Journal of  
1269 Linguistics*, 22, 181–208.

1270 van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When Heuristics Clash  
1271 with Parsing Routines: ERP Evidence for Conflict Monitoring in Sentence  
1272 Perception. *Journal of Cognitive Neuroscience*, 18(7), 1181–1197. doi:  
1273 <https://doi.org/10.1162/jocn.2006.18.7.1181>

1274 van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600  
1275 effects elicited by semantic anomalies. *Cognitive Brain Research*, 22(2),  
1276 241–255. doi: <https://doi.org/10.1016/j.cogbrainres.2004.09.002>

1277 Van Petten, C., & Luka, B. J. (2012). Prediction during lan-  
1278 guage comprehension: Benefits, costs, and ERP components. *In-  
1279 ternational Journal of Psychophysiology*, 83(2), 176–190. doi:  
1280 <https://doi.org/10.1016/j.ijpsycho.2011.09.015>

1281 Xu, X., & Zhou, X. (2016). Topic shift impairs pronoun resolution during  
1282 sentence comprehension: Evidence from event-related potentials. *Psy-  
1283 chophysiology*, 53(2), 129–142. doi: <https://doi.org/10.1111/psyp.12573>

<sup>1286</sup> Zehr, J., & Schwarz, F. (2018). *PennController for Internet Based Experiments*  
<sup>1287</sup> (*IBEX*). doi: <https://doi.org/10.17605/OSF.IO/MD832>

1288     **A    Stimuli**

- 1289     1. Ein Tourist wollte seinen riesigen Koffer mit in das Flugzeug nehmen. Der Koffer war allerdings so schwer, dass die Dame am Check-in entschied, dem Touristen eine extra Gebühr zu berechnen. Daraufhin öffnete der Tourist seinen Koffer und warf einige Sachen hinaus. Somit wog der Koffer des einfallsreichen Touristen weniger als das Maximum von 30 Kilogramm.  
1290     Dann [verabschiedete / wog / unterschrieb] die Dame den Touristen und danach ging er zum  
1291     Gate.
- 1292     2. Ein engagierter Lehrer sah eine alte Weltkarte in der Vitrine eines Antiquitätengeschäfts.  
1293     Ein solch authentisches Artefakt schien dem Lehrer sehr geeignet für sein Klassenzimmer  
1294     zu sein und er sprach die Verkäuferin an. Aufgereggt fragte der Lehrer die sympathische  
1295     Verkäuferin, wie viel die Weltkarte kosten sollte. Obwohl er für eine zusätzliche Weltkarte  
1296     selbst bezahlen musste, sagte der Lehrer der Verkäuferin, dass er dies gerne tun würde. Die  
1297     Verkäuferin sagte daraufhin, wie beschämend es sei, dass die Schule nicht einmal für eine  
1298     Weltkarte bezahlen würde.  
1299     Dann [kaufte / küsste / füllte] der Lehrer die Weltkarte und danach verließ er das Geschäft.
- 1300     3. Eine Redakteurin hatte von ihrer Firma eine Streifenkarte erhalten. Mit dieser Streifenkarte  
1301     konnte die Redakteurin günstig mit dem Bus zur Arbeit fahren und musste nicht jedes Mal  
1302     eine Karte bei dem Busfahrer kaufen. Leider hatte die Tochter der Redakteurin eines Tages  
1303     eine Zeichnung auf die Streifenkarte gemalt. Deswegen hatte die Redakteurin etwas Angst,  
1304     als sie bemerkte, dass der Busfahrer heute nicht gut gelaunt war, als sie ihm die Streifenkarte  
1305     überreichte.  
1306     Dann [stempelte / beschimpfte / aß] der Busfahrer die Streifenkarte und sofort fuhr er viel  
1307     zu schnell weiter.
- 1308     4. Während er einen Tisch baute, brach ein Schreiner seinen schönen Hammer in zwei Teile.  
1309     Der Schreiner hatte den Hammer immer gemocht. Deswegen schien es ihm eine Schande,  
1310     ihn einfach wegzwerfen. Es erschien dem Schreiner eine viel bessere Idee, den Hammer von  
1311     seinem Lehrling reparieren zu lassen.  
1312     Dann [nahm / belächelte / aß] der Lehrling den Hammer und sofort machte er sich an die  
1313     Arbeit.
- 1314     5. Ein Opa wollte einen Apfelkuchen bei einem Konditor kaufen. Der Konditor versicherte  
1315     dem Opa, dass der Apfelkuchen heute besonders gelungen sei. Der Opa schaute auf den  
1316     Apfelkuchen in der Vitrine und sah glücklich den Konditor an.  
1317     Daraufhin [verpackte / belächelte / spülte] der Konditor den Apfelkuchen und dann wandte  
1318     er sich an den nächsten Kunden.
- 1319     6. Eine Lieferbotin brachte einem nervigen Kunden eine Frühlingsrolle. Der Kunde forderte  
1320     jedoch von der Lieferbotin eine neue Frühlingsrolle, da diese kalt war. Nach einer Stunde  
1321     kehrte die Lieferbotin einfach mit derselben kalten Frühlingsrolle zum Kunden zurück.  
1322     Nichtsahnend [nahm / begrüßte / reparierte] der Kunde die Frühlingsrolle und sogleich  
1323     schloss er hinter sich die Tür.
- 1324     7. In einem Restaurant unterhielt sich eine Vegetarierin mit einem befreundeten Metzger über  
1325     eine Fleischwurst auf seinem Teller. Der Metzger sah die Vegetarierin an und erklärte, diese  
1326     Fleischwurst zu essen, wäre ein reines Vergnügen. Er verglich es sogar damit, eine schöne  
1327     Oper zu hören. Die Vegetarierin hielt dies jedoch für einen schlechten Vergleich und wies  
1328     den Metzger darauf hin, dass ein Tier für diese Fleischwurst getötet worden war.  
1329     Dann [durchschnitt / belächelte / mietete] der Metzger die Fleischwurst und sofort begann  
1330     er zu essen.
- 1331     8. Ein gemeiner Kutscher schlug seinen Gaul immer sehr heftig mit einer Peitsche. Eines Tages  
1332     wurde der Kutscher dabei von einem Tierliebhaber beobachtet, der Mitleid mit dem Gaul  
1333     hatte.

- 1336 hatte. Sofort lief der Tierliebhaber zum Kutscher und seinem Gaul und nahm ihm die  
1337 Peitsche weg.  
1338 Dann [bedrohte / streichelte / füllte] der Tierliebhaber den Kutscher und darüber hinaus  
1339 forderte er ihn auf, den Gaul in Ruhe zu lassen.
- 1340 9. Mitten im Meer sah ein Kapitän ein Pärchen auf einem kleinen Segelboot. Schon aus großer  
1341 Entfernung konnte der Kapitän sehen, dass das Segelboot kaputt und das Pärchen in großer  
1342 Not war. Schnell änderte der Kapitän seinen Kurs und steuerte zum Segelboot, um dem  
1343 Pärchen zu helfen.  
1344 Dann [bestieg / rettete / verschloss] der Kapitän das Segelboot und sofort half er dem  
1345 Pärchen.
- 1346 10. Da der Wasserhahn einer älteren Hausfrau nicht mehr aufhörte zu tropfen, rief die Haus-  
1347 frau schließlich einen Handwerker. Zuerst betrachtete der Handwerker den Wasserhahn  
1348 ausführlich und versuchte dann, ihn zu reparieren. Geduldig wartete die Hausfrau daneben.  
1349 Nach einer Weile sagte der Handwerker, dass der Wasserhahn schon zu kaputt sei und er  
1350 einen neuen installieren müsse.  
1351 Daraufhin [lobte / ersetzte / knickte] die Hausfrau den Handwerker und noch lange ärgerte  
1352 sie sich über die Mängel moderner Geräte.
- 1353 11. In einer fremden Stadt buchte ein Urlauber eine Stadtführung. Der Guide freute sich über  
1354 das Interesse des Urlaubers und schenkte ihm noch einen Flyer. Der Guide erklärte dem  
1355 verwunderten Urlauber, dass der Flyer zusätzliche Informationen enthalte, auf die er selbst  
1356 während der Führung nicht eingehen werde. Der Urlauber freute sich über den Flyer und  
1357 dankte dem Guide.  
1358 Nach der Führung [faltete / lobte / kochte] der Urlauber den Flyer und dann machte er sich  
1359 auf den Weg zu seinem Hotel.
- 1360 12. Ein Paparazzi stellte seine große Kamera auf und wartete auf eine berühmte Schauspielerin.  
1361 Es war eine sehr gute Kamera und er wollte unbedingt tolle Bilder schießen. Als die Schaus-  
1362 pielerin den Paparazzi entdeckte, wurde sie sehr wütend, da sie nicht fotografiert werden  
1363 wollte. Deshalb warf die Schauspielerin die Kamera um.  
1364 Daraufhin [bedrohte / schulterte / färbte] der Paparazzi die Schauspielerin und ferner sagte  
1365 er, dass er sich so nicht behandeln lasse.
- 1366 13. Ein Schneider und seine Assistentin suchten für eine neue Schaufensterpuppe, die der Schnei-  
1367 der auf einer Messe ersteigert hatte, einen Platz in dem Laden. Zuerst stellte die Assistentin  
1368 sie in den hinteren Teil des Ladens. Doch dann überzeugte sie den Schneider, die Schaufen-  
1369 sterpuppe in die Nähe des Eingangs zu stellen, da das Licht dort besser war. Tatsächlich  
1370 befand die Assistentin, dass die Schaufensterpuppe dort durch das viele Licht sehr gut zur  
1371 Geltung komme.  
1372 Daraufhin [lobte / bewunderte / schnitt] der Schneider die Assistentin und dann sagte er,  
1373 dass der Platz am Eingang eine gute Idee war.
- 1374 14. Ein Schwimmer übte einen besonders schwierigen Sprung vom Sprungbrett, als er am Beck-  
1375 enrand ein Mädchen entdeckte. Seit einiger Zeit schon bewunderte er das Mädchen aus  
1376 der Ferne, hatte sich aber nie getraut, es anzusprechen. Doch heute wollte der Schwimmer  
1377 dies nachholen und ihm kam die Idee, dass er es mit dem anspruchsvollen Sprung vom Brett  
1378 beeindrucken könnte. So wartete er einen Moment ab, in dem das Mädchen zum Brett blickte  
1379 und sprang dann ins Wasser. Nach dem geglückten Sprung ging der Schwimmer sofort zu  
1380 dem Mädchen und sprach es an.  
1381 Danach [musterte / bewertete / salzte] das Mädchen den Schwimmer und nach einer Weile  
1382 verriet es ihm seine Handynummer.
- 1383 15. Erfreut zeigte eine Sekretärin ihrem Chefarzt die neue Diktiermaschine. Damit konnte der  
1384 Chefarzt seine Arztberichte nun selbst aufzeichnen und war nicht mehr auf die Hilfe seiner

- 1385 Sekretärin angewiesen. Bisher hatte sie nämlich seine Berichte selbst aufschreiben müssen.  
1386 Deswegen freute sie sich besonders über die neue Diktiermaschine. Da der Chefarzt heute  
1387 besonders viele Patienten gehabt hatte, schlug die Sekretärin ihm vor, die neue Diktiermas-  
1388 chine direkt auszuprobieren.  
1389 Dann [verabschiedete / enthüllte / leerte] der Chefarzt die Sekretärin und dann machte er  
1390 Feierabend.
- 1391 16. Eine Reporterin wollte einen Bericht über eine Farm schreiben. Dafür hatte sie sich ein  
1392 paar Fragen überlegt, die sie dem Bauern stellen wollte. Am Hof angekommen begrüßte ein  
1393 Mitarbeiter die Reporterin freundlich und brachte sie zum Bauern. Auf dem Weg erzählte  
1394 der Mitarbeiter, dass er schon seit zwanzig Jahren auf der Farm arbeite. Beim Farmhaus  
1395 angekommen, stellte der Mitarbeiter die Reporterin dem Bauern vor und wünschte ihnen ein  
1396 erfolgreiches Interview.  
1397 Daraufhin [verabschiedete / befragte / ordnete] die Reporterin den Mitarbeiter und an-  
1398 schließend machte sie ein paar Fotos vom Bauernhof.
- 1399 17. Ein Gärtner war sehr stolz auf seinen schönen neuen Rasenmäher, denn der Rasenmäher  
1400 war so groß, dass man auf diesem sitzen und wie mit einem Auto herumfahren konnte. Das  
1401 erzählte der Gärtner auch der kleinen Tochter seines Chefs. Begeistert fragte die Tochter  
1402 des Chefs, ob sie auch mal fahren dürfe. Die Tochter kletterte neben den Gärtner auf den  
1403 Sitz des Rasenmähers und sie drehten eine große Runde über die Wiese.  
1404 Danach [parkte / verabschiedete / halbierte] die Tochter den Rasenmäher und dann sagte  
1405 sie begeistert, dass sie morgen wiederkommen würde.
- 1406 18. Eine junge Dame wollte einen Edelstein von einem Juwelier beurteilen lassen. Stolz erzählte  
1407 sie ihm, dass sie ihn von ihrer Großtante geerbt habe. Nun wollte die Dame von dem Juwelier  
1408 wissen, um welche Art Edelstein es sich handelte. Der Juwelier betrachtete den Edelstein  
1409 sehr lange und sagte dann zu der jungen Dame, dass er sehr selten und wunderschön sei.  
1410 Entzückt [entlohnnte / bestaunte / würzte] die Dame den Juwelier und danach bedankte sie  
1411 sich für sein Fachwissen.
- 1412 19. Ein Mechaniker machte einige Zaubertricks mit einem Schraubenzieher für seine kleine  
1413 Nichte. Zu ihrer Überraschung war das Werkzeug plötzlich aus der Hand des Mechanikers  
1414 verschwunden, doch kurz darauf zog er den Schraubenzieher hinter dem Ohr der Nichte her-  
1415 vor und lachte über ihren erstaunten Gesichtsausdruck. Geheimnisvoll erzählte der Mechaniker  
1416 der Nichte, dass er gerade Magie benutzt habe, um den Schraubenzieher verschwinden zu  
1417 lassen.  
1418 Verblüfft [nahm / bewunderte / kochte] die Nichte den Schraubenzieher und dann sagte sie,  
1419 dass sie noch mehr Zaubertricks sehen wolle.
- 1420 20. Ein Mopedfahrer war versehentlich gegen die Stoßstange eines Autos gefahren. Der Auto-  
1421 fahrer verlangte nun, dass der Mopedfahrer für den Schaden aufkomme, doch dieser weigerte  
1422 sich und sagte, dass die Stange ja überhaupt nicht beschädigt sei. Daraufhin rief der Auto-  
1423 fahrer einen Polizisten zur Hilfe. Der Polizist eilte sofort herbei und begutachtete das  
1424 Fahrzeug. Dann sagte der Polizist zum Mopedfahrer, dass dieser für die Reparaturkosten  
1425 des Autofahrers aufkommen müsse.  
1426 Daraufhin [bestach / entschädigte / sortierte] der Mopedfahrer den Polizisten und außerdem  
1427 entschuldigte er sich für den Unfall.
- 1428 21. Ein Segler und seine Freundin hatten einen Bootsausflug gemacht. Nun wollten sie das Boot  
1429 wieder am Steg festbinden. Die Freundin griff nach dem Strick und wollte dem Segler helfen,  
1430 doch dieser sagte der Freundin, dass er keine Hilfe benötige. Daraufhin packte er den Strick  
1431 und wollte einen Knoten binden. Plötzlich glitt dem Segler der Strick aus den Händen und  
1432 fiel ins Wasser.  
1433 Daraufhin [ermahnte / schnappte / verschraubte] die Freundin den Segler und dann sagte  
1434 sie, er solle etwas aufmerksamer sein.

- 1435        22. Als Piraten von riesigen Goldschätzen auf einer kleinen Insel mitten im Meer gehört hatten,  
1436        machten sie sich sofort auf den Weg, um sie zu suchen. Zu ihrer Überraschung entdeckten sie  
1437        Einheimische auf der Insel, die die Goldschätze bewachten. Die Piraten versteckten sich vor  
1438        den Einheimischen, um in Ruhe ihren Überfall vorbereiten zu können. Die Piraten warteten  
1439        ab, bis die Einheimischen schliefen, um unbemerkt an die Goldschätze zu kommen.  
1440        Dann [versklavten / raubten / wechselten] die Piraten die Einheimischen und danach segelten  
1441        sie Richtung Heimat.
- 1442        23. Ein Junge verspürte Lust, einen Apfel zu essen. Erst gestern hatte er bei der Ernte geholfen  
1443        und anschließend den vollen Korb nach Hause getragen. Bei dem Gedanken, wie schwer der  
1444        Korb gewesen war und daran, wie frisch und saftig der Apfel sein musste, lief dem Jungen  
1445        glatt das Wasser im Mund zusammen. Der Junge wusste, dass die Mutter den Korb mit  
1446        seinem ersehnten Apfel im Keller versteckte.  
1447        Sofort [suchte / zerschnitt / schlug] der Junge den Korb und dann entschied er sich für einen  
1448        großen roten Apfel.
- 1449        24. Schon seit einiger Zeit bereitete sich ein Sportler auf einen großen Wettkampf im Ringen  
1450        vor. Der Vater des Sportlers half ihm täglich beim Training, denn gemeinsam wollten sie  
1451        den Juror mit einer guten Technik überzeugen. Der Vater kannte den Juror schon seit  
1452        langer Zeit und wusste, dass der Juror sehr auf die richtige Technik achtete. Am Tag des  
1453        Wettkampfes war der Vater sehr aufgereggt, doch der Sportler beeindruckte alle mit seiner  
1454        hervorragenden Technik und gewann den Wettbewerb.  
1455        Danach [beglückwünschte / entdeckte / öffnete] der Juror den Sportler und außerdem lobte  
1456        er dessen Sohn in höchsten Tönen.
- 1457        25. Eine Geschäftsfrau hatte bei einer Auktion eine süße, alte Scheune ersteigert, die sie zu  
1458        einer Bar herrichten ließ. Ihr Mann war nämlich Kellner und wollte sich schon lange selb-  
1459        stständig machen. Da der Mann in Bezug auf Ästhetik nicht sehr viel verstand, überließ  
1460        er es ihr, die Renovierungsarbeiten anzuleiten. Diese hatten einige Zeit beansprucht, doch  
1461        der Geschäftsfrau war das egal, denn sie war mit dem Resultat äußerst zufrieden. Die Bar  
1462        war wunderschön geworden und hatte ihr altes Flair nicht verloren. Begeistert zeigte die  
1463        Geschäftsfrau ihrem Mann die fertige Bar.  
1464        Daraufhin [umarmte / bestaunte / sortierte] der Mann die Geschäftsfrau und dann lobte er  
1465        sie für ihren guten Geschmack.
- 1466        26. Ein Rentner wollte auf einem Trödelmarkt sein altes Zelt verkaufen, mit dem er schon viele  
1467        schöne Urlaube verbracht hatte. Deshalb wollte er nun einen neuen Besitzer finden, der  
1468        genauso viel Freude daran haben würde, wie er selbst sie gehabt hatte. Plötzlich tauchte  
1469        ein kleines Kind neben ihm auf und starrte begeistert auf das Zelt. Das Kind stellte dem  
1470        Rentner viele Fragen und erzählte ihm auch von seinen eigenen Campingausflügen mit der  
1471        Familie. Schließlich fragte das Kind nach dem Preis für das Zelt.  
1472        Lachend [holte / tätschelte / aß] der Rentner das Zelt und dann schenkte er es dem Kind.
- 1473        27. Als ein Lehrer seine Unterlagen holen wollte, bemerkte er, dass er seine Tasche nicht bei  
1474        sich hatte. Erschrocken überlegte er, wo er die Tasche hatte stehen lassen. Ihm fiel ein,  
1475        dass er sich eine Limonade hatte kaufen wollen, aber nicht genügend Kleingeld gehabt hatte.  
1476        Deswegen war der Lehrer nochmal zurück ins Lehrzimmer gegangen, um mehr Geld zu  
1477        holen. Dort war er von einem Kollegen in ein wichtiges Gespräch verwickelt worden, sodass er  
1478        die Limonade total vergessen hatte. In aller Aufregung über die Limonade hatte er bestimmt  
1479        auch die Tasche in der Kantine stehen lassen.  
1480        Zurück in der Kantine [kaufte / fand / unterrichtete] der Lehrer die Limonade und dann  
1481        suchte er seine Tasche.
- 1482        28. Eine junge Bergsteigerin hatte eine neue Spitzhacke geschenkt bekommen und war nun er-  
1483        picht darauf, diese sogleich an einer sehr steilen Bergwand auszuprobieren. Ihre Mutter

- 1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533
- hatte ihr die Spitzhacke erst am Tag zuvor gekauft, nachdem der Verkäufer der Mutter versichert hatte, dass es ein sehr gutes Modell sei. Am Morgen hatte die Mutter ihr viel Erfolg gewünscht und danach war die Bergsteigerin voller Tatendrang aufgebrochen, den Berg zu erklimmen. Doch leider brach die Spitzhacke durch, nachdem die Bergsteigerin schon eine Weile geklettert war und sie musste von der Bergwacht gerettet werden.  
Im Krankenhaus [tröstete / verwünschte / stapelte] die Mutter die Bergsteigerin und hinterher betrachtete sie die kaputte Spitzhacke.
29. Ein Förster und eine Praktikantin gingen in den Wald, um Wild zu sehen. Der Förster schlug vor, auf einen Hochsitz zu klettern, da sie dort einen besseren Überblick haben würden. Nach einer Weile entdeckte die Praktikantin einen Hirsch. Der Hirsch war groß und hatte ein mächtiges Geweih. Doch er war sehr weit entfernt, weshalb die Praktikantin enttäuscht sagte, dass sie kaum etwas erkennen könne. Daraufhin holte der Förster ein Fernglas aus seiner Tasche und gab es ihr, damit sie den Hirsch sehen konnte.  
Dann [umarmte / beobachtete / sammelte] die Praktikantin den Förster und anschließend bedankte sie sich für das Fernglas.
30. Ein Angeklagter wurde zum Gerichtssaal gebracht, wo der Richter und der Staatsanwalt schon auf ihn warteten. Der Mann wurde eines Raubüberfalls beschuldigt und heute war der erste Anhörungstag. Nachdem der Richter die Sitzung eröffnet hatte, trug der Staatsanwalt alle Punkte vor, die dem Angeklagten vorgeworfen wurden. Danach dankte der Richter dem Staatsanwalt und begann mit der Anhörung des Angeklagten.  
Am Ende [konsultierte / befragte / kopierte] der Richter den Staatsanwalt und danach ließ er den ersten Zeugen herein.
31. Aufgeregzt standen die Gäste in der Kirche und lauschten der röhrenden Predigt des Pfarrers. Die Braut konnte den Moment kaum erwarten, in dem sie dem Bräutigam ihr Jawort geben und den Ring erhalten würde. Sie wusste, dass der Ring ein sehr besonderes Erbstück aus der Familie des Bräutigams war, das schon lange von Generation zu Generation weitergegeben worden war, und fühlte sich sehr geehrt, dieses zu erhalten. Als der Pfarrer die Predigt beendete und dem Brautpaar die Frage stellte, gaben sich der Bräutigam und die Braut das Jawort, während der Trauzeuge den schönen Ring hervorholte.  
Glücklich [küsst / bewunderte / vereinfachte] die Braut den Bräutigam und dann übergab der Trauzeuge den Ring.
32. Während ein Ritter seinen Umhang anprobieren, besprach er das bevorstehende Turnier mit dem Burgfräulein. Das Burgfräulein fand den Umhang viel zu groß und schlug vor, ihn etwas zu kürzen. Aber der Ritter wollte nicht, dass das Burgfräulein irgend etwas veränderte. Er hatte den Umhang schon seit Jahren und dieser hatte dem Ritter bisher immer Glück gebracht.  
Daraufhin [faltete / verspottete / entleerte] das Burgfräulein den Umhang und dann wünschte es dem Ritter viel Erfolg.
33. In einem Museum konnte eine Besucherin einen bestimmten Raum nicht finden. Verzweifelt versuchte sie, sich an der Wegbeschreibung auf ihrer Eintrittskarte zu orientieren, aber ohne Erfolg. Dann entdeckte die Besucherin eine Aufsichtsperson am anderen Ende des Raumes und fragte sie nach Hilfe. Die Aufsichtsperson erzählte, dass einige Leute Probleme mit der Wegbeschreibung auf der Eintrittskarte hätten. Die Aufsichtsperson nahm die Eintrittskarte der Besucherin und versprach, ihr den Weg zu zeigen.  
Daraufhin [begleitete / studierte / erfand] die Aufsichtsperson die Besucherin und währenddessen erklärte sie ihr den Weg.
34. Ein Händler war auf dem Weg in den fernen Orient, um dort kostbare Gewürze einzukaufen. Dort angekommen begab er sich zum Marktplatz. Der Händler konnte schon von weitem die Rufe hören, mit denen die Sklaven zum Kauf angepriesen wurden. Der Händler fragte jemanden nach dem Stand mit den Gewürzen. Auf dem Weg zu den Gewürzen kam auch er

- 1534 an den Sklaven vorbei, welche seine fremdländischen Gewänder interessiert musterten.  
1535 Dann [grüßte / kaufte / versiegelte] der Händler die Sklaven und anschließend ging er weiter  
1536 zu den Gewürzen
- 1537 35. Ein Kind entdeckte in einem Schaufenster einen Teddybären, den es unbedingt haben wollte.  
1538 Der Ladenbesitzer bemerkte die bewundernden Blicke des Kindes und nahm ihn vom Regal.  
1539 Das Kind sagte dem Ladenbesitzer, dass es den Teddybären gerne kaufen würde, worauf der  
1540 Ladenbesitzer ihm den Teddybären überreichte.  
1541 Dann [drückte / entlohnnte / bastelte] das Kind den Teddybären und dann lachte es vor  
1542 Freude.
- 1543 36. Eine Hundeliebhaberin hatte ihren Nachbarn engagiert, um auf den Welpen aufzupassen,  
1544 da sie über das Wochenende geschäftlich unterwegs war. Da die Hundeliebhaberin wusste,  
1545 dass der Nachbar sich gut mit Tieren auskannte und den Welpen auch sehr gerne hatte,  
1546 hatte sie keine Bedenken. Trotzdem war sie froh, als sie wieder zu Hause war. Als die  
1547 Hundeliebhaberin die Haustüre aufschloss, rannte ihr der Welpe entgegen und der Nachbar  
1548 begrüßte sie freundlich.  
1549 Daraufhin [entlohnnte / drückte / sortierte] die Hundeliebhaberin den Nachbarn und außerdem  
1550 bedankte sie sich für seine Zeit.
- 1551 37. Ein Schuhverkäufer hatte gerade einem Kunden ein Paar Schuhe verkauft, als er beobachtete,  
1552 wie draußen vor seinem Laden ein Dieb dem Kunden seine Geldbörse entwendete. Auch sah  
1553 der Schuhverkäufer, dass dieser nichts davon mitbekommen hatte und der Dieb sich geschickt  
1554 aus dem Staub machte. Der Schuhverkäufer blickte dem Kunden hinterher und rannte schnell  
1555 nach draußen, um den Dieb aufzuhalten.  
1556 Dann [bemitleidete / verfolgte / hinterlegte] der Schuhverkäufer den Kunden und sofort  
1557 erzählte er ihm von dem beobachteten Diebstahl.
- 1558 38. Ein Eskimo wollte auf die Jagd gehen, um eine Robbe zu jagen. Er nahm seine Freundin  
1559 als Begleitung mit. Auf dem Weg sagte der Eskimo zu der Freundin, dass sie sich ganz  
1560 still verhalten müsse und sich nicht mehr bewegen dürfe, sobald sie die Robbe erblickten.  
1561 Nach einer Weile entdeckte der Eskimo die Robbe in geeigneter Entfernung und zeigte sie  
1562 der Freundin.  
1563 Dann [ermahnte / erschoss / verpackte] der Eskimo die Freundin und danach lud er sein  
1564 Gewehr neu.
- 1565 39. Nach einer Abendveranstaltung machte sich eine Tänzerin auf den Weg nach Hause. Sie  
1566 beeilte sich, um schnell bei ihrer Tochter und der Babysitterin zu sein. Da die Babysitterin  
1567 das erste Mal auf die Tochter aufgepasst hatte, wollte die Tänzerin schnell nach Hause,  
1568 um nach dem Rechten zu schauen. Zuhause angekommen fand die Tänzerin eine glückliche  
1569 Tochter und eine entspannte Babysitterin vor und war sehr erleichtert.  
1570 Dann [vergütete / umarmte / stapelte] die Tänzerin die Babysitterin und anschließend  
1571 schickte sie diese nach Hause.
- 1572 40. Ein Minister und sein Berater waren erzürnt über den Präsidenten aus dem Nachbarland, da  
1573 dieser sich nicht an ein Handelsabkommen hielt. Daraufhin riet der Berater dem Minister,  
1574 mit Sanktionen gegen den Präsidenten vorzugehen. Der Berater organisierte ein Treffen, bei  
1575 dem der Minister dem Präsidenten seine Forderungen überbringen konnte.  
1576 Dann [verhandelte / feierte / schminkte sich] der Minister mit dem Präsidenten und dabei  
1577 besprachen sie genauere
- 1578 41. Seit Monaten hatte sich der Athlet mit der Trainerin darauf vorbereitet, bei dem wichtigsten  
1579 Wettkampf des Jahres den Pokal zu holen. Die Trainerin trieb ihn hart an, da sie sicher war,  
1580 dass er gute Chancen hatte. Und tatsächlich hatte sich die harte Arbeit gelohnt, denn der  
1581 Athlet gewann den Pokal und überglücklich bedankte er sich bei der Trainerin. Stolz hielt  
1582 der Athlet den Pokal in den Händen.

- 1583        Im Hotel [polierte / bejubelte / verspeiste] die Trainerin den Pokal und anschließend stellte  
1584        sie ihn auf den Tisch.
- 1585        42. Ein Autor ging mit dem Hund spazieren, um an der frischen Luft neue Ideen für sein derzeitige  
1586        ges Buch zu bekommen. Der Autor hatte einen Ball dabei, da der Hund sehr verspielt war.  
1587        Im Park angekommen, warf der Autor den Ball einige Meter weit. Sofort rannte der Hund  
1588        dem Ball nach und brachte ihn brav zurück.  
1589        Daraufhin [nahm / tätschelte / zitierte] der Autor den Ball und wieder warf er ihn einige  
1590        Meter weit.
- 1591        43. Eine Oma und ein Kleinkind standen vor einem Hasenstall und streichelten das Kaninchen.  
1592        Die Oma gab dem Kleinkind Löwenzahn, damit dieses das Kaninchen füttern konnte und  
1593        dann ging sie noch mehr Löwenzahn holen. Doch plötzlich biss das Kaninchen das Kleinkind  
1594        und dieses fing fürchterlich an zu weinen.  
1595        Daraufhin [fütterte / streichelte / strickte] die Oma das Kaninchen und nebenbei tröstete  
1596        sie das Kleinkind.
- 1597        44. Der Geschäftsführer und der Coach saßen nebeneinander und schauten einem bedeutenden  
1598        Fußballspiel zu. Leider war die Mannschaft, die der Coach trainierte, deutlich unterlegen.  
1599        Der Torwart hatte bisher fast keinen Ball gehalten. Der Geschäftsführer saß bekümmert auf  
1600        der Bank und selbst die gute Leistung der anderen Spieler konnte die Uneschicktheit des  
1601        Torwarts nicht wieder gut machen. Auch der Coach wirkte verzweifelt, als der Torwart aus  
1602        Versehen den Ball einem gegnerischen Spieler zuspielte, worauf dieser ein Tor schoss. Am  
1603        Ende verlor die Mannschaft das Spiel. Entrüstet sagte der Geschäftsführer dem traurigen  
1604        Coach, dass er mit dem Torwart reden wolle.  
1605        Letztendlich [suspendierte / umarmte / reparierte] der Geschäftsführer den Torwart und  
1606        dann fuhr er immer noch wütend nach Hause.
- 1607        45. Als ein Referendar den Weihnachtsmarkt seines Gymnasiums betrat, wurde er direkt von  
1608        ein paar Schülern begrüßt. Die Schüler berichteten dem Referendar, dass sie eine Tombola  
1609        organisiert hatten und nun versuchten, die Lose zu verkaufen. Die Schüler hatten schon sehr  
1610        viele Lose verkauft und erzählten dem Referendar nun, was er alles Schönes mit den Losen  
1611        gewinnen könne.  
1612        Amüsiert [kaufte / musterte / betrat] der Referendar die Lose und tatsächlich gewann er  
1613        einen Preis.
- 1614        46. Eine Mutter ging mit ihren eineiigen Zwillingen zum Doktor, da diese geimpft werden sollten.  
1615        Im Behandlungszimmer des Doktors machte dieser Witze darüber, wie ähnlich sich die  
1616        Zwillinge sahen und zeigte ihnen die Spritzen, die er schon vorbereitet hatte. Der Doktor  
1617        versicherte ihnen, dass sie keine Angst vor den Spritzen haben müssten. Da die Spritzen mit  
1618        ihren langen, dünnen Nadeln tatsächlich angsteinflößend aussahen, bekamen die Zwillinge  
1619        trotzdem Angst.  
1620        Dann [nahm / verwechselte / bastelte] der Doktor die Spritzen und anschließend begann er  
1621        mit der Impfung.
- 1622        47. In einem Kriegsgebiet wollte ein Soldat eine Zivilistin unbemerkt an den gegnerischen Truppen  
1623        vorbei schmuggeln, da es für sie sehr gefährlich war, allein unterwegs zu sein. Da sie sich  
1624        in einem Kriegsgebiet befanden, hielt der Soldat die Waffe bereit. So schllichen die Zivilistin  
1625        und der Soldat mit seiner Waffe still die Häuser entlang. Die Zivilistin war sehr erleichtert  
1626        über die Hilfe und fühlte sich durch die Waffe auch sicher, doch plötzlich tauchte vor ihnen  
1627        ein Panzer des gegnerischen Lagers auf.  
1628        Schnell [zückte / versteckte / durchkämmte] der Soldat die Waffe und sofort ging er in  
1629        Deckung.
- 1630        48. Ein Dirigent hatte ein neues Stück geschrieben und wollte es heute Abend zum ersten Mal  
1631        dem Publikum zeigen. Lange hatte er mit dem Orchester geprobt und war gespannt auf die

- 1632 Reaktion des Publikums. Machte das Orchester heute Abend keinen Fehler, könnte das Stück  
1633 die Karriere des Dirigenten voranbringen. Als der Abend gekommen war, betrat der Dirigent  
1634 zusammen mit dem Orchester die Bühne, um dem Publikum das Stück zu präsentieren.  
1635 An diesem Abend [spielte / verzauberte / engagierte] das Orchester das Stück und das  
1636 Publikum applaudierte.
- 1637 49. Ein Doktorand hatte nach Jahren endlich seine Arbeit beendet und musste sie nun seiner  
1638 Betreuerin und anderen Prüfern vorstellen. Obwohl der Doktorand eng mit der Betreuerin  
1639 zusammengearbeitet hatte und wusste, dass die Arbeit sehr gut war, war er trotzdem sehr  
1640 nervös. Vor der Prüfung ging der Doktorand noch einmal die wichtigsten Stichpunkte  
1641 bezüglich der Arbeit durch, dann folgte er den anderen Prüfern ins Büro der Betreuerin.  
1642 Dort [begrußte / verteidigte / reparierte] der Doktorand die Betreuerin und dann hielt er  
1643 seinen Vortrag.
- 1644 50. Eine Protestantin wollte nach Israel fliegen, um sich Jerusalem anzuschauen. Sie hatte die  
1645 Reise geplant, seitdem der Pfarrer ihr Bilder von seinem Aufenthalt dort gezeigt hatte. Nun  
1646 war die Reise fertig organisiert und der Abflug rückte immer näher. Doch die Protestantin  
1647 machte sich Sorgen, da es in letzter Zeit vermehrt Unruhen gegeben hatte. So ging sie zu  
1648 dem Pfarrer, um ihn um Rat zu fragen. Sie wollte, dass der Pfarrer ihr versicherte, dass sie  
1649 sich keine Sorgen machen müsse. Dadurch würde sich die Protestantin bezüglich der Reise  
1650 sicherer fühlen.  
1651 Daraufhin [segnete / bewilligte / las] der Pfarrer die Protestantin und dann wünschte er der  
1652 Protestantin einen guten Flug.
- 1653 51. Eine Erzieherin suchte einen Therapeuten auf. Dieser war ihr von einer Freundin empfohlen  
1654 worden, nachdem sie über Symptome geklagt hatte. Die Symptome waren denen einer De-  
1655 pression ziemlich ähnlich und die Erzieherin hatte beschlossen, dass sie professionelle Hilfe  
1656 von dem Therapeuten brauche. So war die Erzieherin sehr erleichtert gewesen, als sie endlich  
1657 einen Termin bei dem Therapeuten bekommen hatte, da sich die Symptome in letzter Zeit  
1658 noch verschlimmert hatten.  
1659 In der Praxis [erfragte / behandelte / tauschte] der Therapeut die Symptome und daraufhin  
1660 verschrieb er ein Medikament.
- 1661 52. Eine Designerin hatte den Auftrag bekommen, ein Buch grafisch zu gestalten. In dem Buch  
1662 ging es um Geschichten über Eisbären. Die Geschichten waren für Kinder gedacht und der  
1663 Verlag wollte, dass die Designerin die Eisbären bildlich darstellte. Nun hatte die Designerin  
1664 die Geschichten über die Eisbären zu Ende gelesen und war bereit, mit der Arbeit zu beginnen.  
1665 Dann [malte / veranschaulichte / leerte] die Designerin die Eisbären und bis spät in die  
1666 Nacht arbeitete sie an der Geschichte.
- 1668 53. Eines Abends wurde ein Architekt von dem Bürgermeister angerufen. Dieser sagte, dass  
1669 die Stadt eine neue Turnhalle zu bauen beabsichtigte. Er beauftragte den Architekten,  
1670 einen Plan der Turnhalle zu erstellen und diesen in einer Rede vor dem Gemeinderat näher  
1671 auszuführen. In der Rede solle er auf die besonderen Merkmale seines Entwurfes eingehen.  
1672 Der Architekt versicherte, dass er sofort mit der Konzeption der Turnhalle beginnen werde  
1673 und bedankte sich für die Tipps bezüglich der Rede.  
1674 Daraufhin [schrieb / entwarf / rief] der Architekt die Rede und dann goss er sich ein Glas  
1675 Wein ein.
- 1676 54. Ein Gitarrist wurde von einer Agentin engagiert, um zusammen mit einer Sängerin auf einer  
1677 Party aufzutreten. Auf dem Weg zur Probe erzählte die Agentin dem Gitarristen, dass  
1678 sie lange nach einem guten Musiker gesucht habe und glaube, dass seine Art zu spielen  
1679 ausgezeichnet mit der Stimme der Sängerin harmonieren würde. Dann holte der Gitarrist  
1680 sein Instrument und die Agentin sagte, dass die Sängerin schon bereit sei und sie direkt mit  
1681 der Probe beginnen könnten.

- 1682            Dann [verabschiedete / traf / kaufte] der Gitarrist die Agentin und dann ging er schnell zur  
1683            Bühne.
- 1684        55. In einem Museum war ein Kurator dabei, eine neue Ausstellung zu gestalten. Da es um  
1685            plastische Kunst ging, hatte sich der Kurator von einer befreundeten Galeristin eine Skulptur  
1686            geliehen. Gerade war die Galeristin eingetroffen und sie überlegten nun gemeinsam, wo die  
1687            Skulptur am Besten zur Geltung kommen würde. Lange suchten sie nach einem geeigneten  
1688            Platz und fanden schließlich einen. Mühevoll installierte der Kurator die Skulptur, während  
1689            die Galeristin Anleitungen gab.  
1690            Danach [umarmte / betrachtete / sammelte] der Kurator die Galeristin und dabei dankte er  
1691            ihr für ihre Hilfe.
- 1692        56. Eine Studentin war mit einer Kommilitonin in einer Kneipe. Da sie danach noch in einem  
1693            Club feiern gehen wollten, beschlossen sie, sich auf der Toilette frisch zu machen. Dann fragte  
1694            die Kommilitonin die Studentin, ob sie ihre Wimperntusche ausleihen dürfe, da sie ihre eigene  
1695            vergessen hatte. Sofort gab die Studentin ihr die Wimperntusche. Die Kommilitonin fragte  
1696            eine Bedienung nach der Toilette und machte sich mit der Wimperntusche in der Hand auf  
1697            den Weg zur Toilette.  
1698            Dann [betrat / benutzte / las] die Kommilitonin die Toilette und anschließend schminkte sie  
1699            sich.
- 1700        57. Ein Verbrecher war auf dem Weg zu einem Haus, wo ein Ermittler wohnte. Dieser untersuchte  
1701            einen Fall, in den der Verbrecher verstrickt war. Deswegen wollte dieser den Ermittler  
1702            aus dem Weg räumen. Am Haus angekommen verschaffte sich der Verbrecher Zutritt. Er  
1703            wusste, dass es in dem Haus einen Schäferhund gab und bedacht achtete er darauf, dass der  
1704            Schäferhund ihn nicht hörte. Bevor er den Ermittler suchte, gab er dem Schäferhund etwas  
1705            zu Essen, um ihn abzulenken.  
1706            Dann [streichelte / erschoss / faltete] der Verbrecher den Schäferhund und danach machte  
1707            er sich auf die Suche nach dem Ermittler.
- 1708        58. Ein Beschuldigter und seine Anwältin betrat den Gerichtssaal, um bei der bevorstehenden  
1709            Anhörung zu beweisen, dass der Beschuldigte die Tat nicht begangen hatte. Der Kläger saß  
1710            schon an seinem Platz und warf den beiden böse Blicke zu. Die Anwältin ging noch ein  
1711            paar ihrer Unterlagen durch, dann begann die Verhandlung. Der Kläger wurde nach vorne  
1712            gebeten und von der Anwältin zur Tat befragt. Der Beschuldigte blickte nervös drein, als  
1713            der Kläger ihn vor aller Augen der Tat bezichtigte.  
1714            Daraufhin [verteidigte / entließ / schwenkte] die Anwältin den Beschuldigten und dann  
1715            wandte sie sich an den Richter.
- 1716        59. Ein Junge ging mit seinem Kumpel zum See, da er schwimmen wollte. Der Kumpel hatte  
1717            seine Angel dabei und erzählte dem Jungen, dass er heute einen Flussbarsch angeln wollte,  
1718            von denen es viele im See gab. Er hatte einen besonderen Köder dabei, mit dem er den  
1719            Flussbarsch anlocken wollte. Der Junge wünschte dem Kumpel viel Glück mit dem Fluss-  
1720            barsch und machte einen Salto ins Wasser.  
1721            Daraufhin [angelte / beobachtete / trocknete] der Kumpel den Flussbarsch und danach ging  
1722            er selbst ins Wasser.
- 1723        60. Eine Schwangere betrat das Untersuchungszimmer einer Gynäkologin und wurde von der  
1724            Gynäkologin freundlich begrüßt. Die Gynäkologin deutete auf die Liege im Zimmer und  
1725            forderte die Schwangere auf, sich dort hinzulegen. Die Liege war etwas hoch eingestellt,  
1726            doch die Schwangere schaffte es, hochzukommen und legte sich auf die Liege.  
1727            Daraufhin [verstellte / untersuchte / verordnete] die Gynäkologin die Liege und dann wandte  
1728            sie sich der Schwangeren zu.