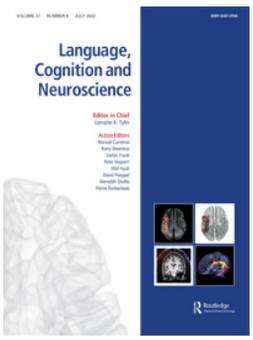


Routledge
Taylor & Francis Group

Language, Cognition and Neuroscience

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/plcp21

Situational expectancy or association? The influence of event knowledge on the N400

Elisabeth Rabs, Francesca Delogu, Heiner Drenhaus & Matthew W. Crocker

To cite this article: Elisabeth Rabs, Francesca Delogu, Heiner Drenhaus & Matthew W. Crocker (2022) Situational expectancy or association? The influence of event knowledge on the N400, Language, Cognition and Neuroscience, 37:6, 766-784, DOI: 10.1080/23273798.2021.2022171

To link to this article: https://doi.org/10.1080/23273798.2021.2022171

LANGUAGE, COGNITION AND NEUROSCIENCE
2022, VOL. 37, NO. 6, 766–784
https://doi.org/10.1080/23273798.2021.2022171

Routledge
Taylor & Francis Group

REGULAR ARTICLE

OPEN ACCESS Check for updates

Situational expectancy or association? The influence of event knowledge on the N400

Elisabeth Rabs, Francesca Delogu, Heiner Drenhaus and Matthew W. Crocker

Department of Language Science & Technology, Saarland University, Saarbrücken, Germany

ABSTRACT

Electrophysiological studies suggest that situational event knowledge plays an important role in language processing, but often fail to distinguish whether observed effects are driven by combinatorial expectations, or simple association with the context. In two ERP experiments, participants read short discourses describing ongoing events. We manipulated the situational expectancy of the target word continuing the event as well as the presence of an associated, but inactive event in the context. In both experiments we find an N400 effect for unexpected compared to expected target words, but this effect is significantly attenuated when the unexpected target is nonetheless associated with non-occurring context events. Our findings demonstrate that the N400 is simultaneously influenced by both simple association with – and combinatorial expectations derived from – situational event knowledge. Thus, experimental investigations and comprehension models of the use of event knowledge must accommodate the role of both expectancy and association in electrophysiological measures.

ARTICLE HISTORY
Received 14 September 2020
Accepted 8 December 2021

KEYWORDS
ERPs; N400; discourse comprehension; event knowledge; association; expectancy

1. Introduction

Over the last few decades, psycholinguistic research has gathered convincing evidence that world knowledge – even though not explicitly stated in the immediate linguistic context – rapidly influences on-line language comprehension. In particular, knowledge about real-world events, so-called *script* or *event knowledge* (Schank & Abelson, 1977) has informed a number of comprehension models and theories (Elman, 2009; Golden & Rumelhart, 1993; Van Dijk & Kintsch, 1983; Venhuizen et al., 2019).

In behavioural studies, the use of event knowledge has been shown to lead to decreased reaction times (Bicknell et al., 2010; Ferretti et al., 2001; Hare et al., 2009; Madden & Zwaan, 2003; McRae et al., 2005, 1998). In the eye-tracking literature, facilitatory effects of event knowledge have been linked to anticipatory eye movements to expected or depicted event participants (Kamide et al., 2003; Knoeferle & Crocker, 2006, 2007; Knoeferle et al., 2005), while event knowledge violations have been linked to longer reading times (Camblin et al., 2007; Kamide et al., 2003).

In ERP research, studies have primarily investigated event knowledge influence on the N400 component, a centro-parietal distributed negative-going wave peaking approximately 400 ms after stimulus onset, which has been shown to be sensitive to semantic processing (for a review, see e.g. Baggio, 2018; Kutas & Federmeier, 2011). Evidence from discourse studies suggests that the N400 is sensitive to what the greater linguistic context is about, with discourse-supported words eliciting attenuated N400s compared to similarly plausible, but unsupported words (Nieuwland et al., 2020; Nieuwland & Van Berkum, 2006; Van Berkum et al., 1999) or violations of world and event knowledge such as *The dutch trains are white* vs. *yellow* (Hagoort 2004; see also Coulson et al., 2005; Metusalem et al., 2012). In particular, the N400 has been shown to be inversely correlated with the expectancy of a word, with lower N400 amplitudes for expected words than for unexpected words (Chwilla & Kolk, 2005; Kutas & Hillyard, 1984; Van Berkum et al., 1999). There has been ongoing debate however whether this expectancy stems from facilitated lexical retrieval of a word from memory (Delogu et al., 2019; Kutas & Federmeier, 2000; Lau et al., 2008), or from facilitated integration of a word into the prior context (Brown & Hagoort, 1993; Van Berkum et al., 2005), or from a hybrid system entailing both (Baggio & Hagoort, 2011; Nieuwland et al., 2020). As word expectancy can be explained under a retrieval account as well as an integration account, it is

CONTACT Elisabeth Rabs erabs@coli.uni-saarland.de
Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/23273798.2021.2022171.

not the primary focus of this study to differentiate between the two. We will return to this point in the general discussion.

In addition, a large number of studies has shown that the N400 is also influenced by semantic association, with words eliciting a smaller N400 when preceded by semantically related words (Holcomb & Neville, 1990; Kutas & Hillyard, 1989; Neely, 1991; Swaab et al., 2002). Effects of semantic association, or priming, have been found not only for isolated words, but also in sentence context (Van Petten, 1993; Van Petten et al., 1997). For studies investigating the use of event knowledge, this poses a problem, as in many cases it is not entirely clear whether the observed N400 modulations reflect facilitation from simple association between words in the context and a target word or from what we will for the rest of this paper refer to as "situational expectancy", that is the expectations about the unfolding situation described in a discourse. Such expectations are often based on our detailed knowledge about events and the world, but they specifically refer to the part of this knowledge that evaluates a given situation as expected or unexpected.

This is different from the general knowledge about likely participants, locations and objects within an event. For example, our general event knowledge makes both "food" and "menu" expected participants in a "going to the restaurant" event, and thus both words are associated with the word "restaurant"; but in a description of this event, after encountering "ordering food", "menu" is unlikely to be mentioned again, and thus situationally unexpected, as "reading the menu" should have happened before ordering food. This distinction is usually not made clear in studies of event knowledge.

To illustrate, consider passages describing events as in (1), taken from Metusalem et al. (2012):

(1) *Going to the movies is great fun. Before the show starts, I like to get a snack. There's nothing like watching the show while eating a big box of popcorn/soda/car.*

They find a reduced N400 on the critical word for contextually expected words like *popcorn* compared to event-related, but incongruent and hence unexpected words like *soda* and the largest N400 effect for event-unrelated and incongruent words like *car*. Findings like these have been taken as evidence that the N400 indexes the rapid use of detailed event knowledge to facilitate processing of words that are expected given the current situation, even if they may be incongruent with the linguistic context, as the smaller N400 for an anomalous but event-related target object like *soda* compared to an event-unrelated target like *car* shows.

There is, however, an alternative explanation: in examples like (1), the notion of detailed event knowledge often confounds the situational expectations arising from event knowledge with the associative knowledge about event components. Instead, these graded N400 effects can often be explained in terms of association: The greater facilitation for an expected word like *popcorn* might as well stem from the fact that the preceding context offers more words that are associated with it, like *movies, show, snack, eating, box*, than with the unexpected target words. For example, *soda* might only be associated with *movies, show, snack* and also not be a suitable thematic role filler for the verb *eating*, while an event-unrelated word like *car* is not associated with any of the words in the context, nor is it a suitable thematic role filler. This interpretation raises the question as to whether a situationally unexpected version of example (1), e.g. ending in "there's nothing like watching the show *before* eating a box of popcorn", would result in a similar N400 reduction. This concern holds for a number of studies involving world or event knowledge: for example, the reduced N400 on *tea* compared to *coffee* in Kutas and Hillyard (1984) (*He liked lemon and sugar in his tea/coffee*) can be explained by association between *lemon*, *sugar* and *tea*. The reduced N400 in Coulson et al. (2005) on *shoes* compared to *oil* in *They were hard to walk in, but she loved her olive shoes/oil* may not only be a sentence congruency effect, but also stem from association with *walk in*. Likewise, the reduced N400 in Bicknell et al. (2010) (*The journalist* vs. *the mechanist checked the spelling*) observed on *spelling* can be attributed to a semantic association between *journalist* and *spelling*.

Association has traditionally been used to describe the relation between single words that semantically prime each other, within or without context, as stated above. The way we use the term in this study, however, entails a somewhat broader definition: There is evidence that association between words and concepts may be formed by our knowledge about events and their participants (Camblin et al., 2007; Chwilla & Kolk, 2005; Ferretti et al., 2001; Hare et al., 2009; McRae et al., 2005). For example, Hare et al. (2009) find that events prime event participants in word pairs with low free association norms (for example, sale – *clothes*, reunion – *friends* and interrogation – *table*). Similarly, Chwilla and Kolk (2005) find a reduced N400 on the last word in word triplets that are only associated via an event and would not prime each other out of context (for example, director - bribe – *dismissal*) compared to unrelated word triplets. These results are

especially interesting as they imply that even in the absence of sentential context, event knowledge may influence word processing, thus further modulating the association between words. Importantly, association in this case may not necessarily be captured by norming studies or computational models of word association and therefore is challenging to quantify. That is, it may entail more than a direct relation between two words that share similarities in meaning, a logical connection or a cooccurrence pattern. It may be mediated by our knowledge about the world, in which words or concepts are connected via events, situations and experiences. Association may extend to more than the immediate linguistic context of a word and be shaped by a comprehender's knowledge about this context. However, studies indicating that words may be associated through event knowledge, which is reflected in N400 amplitude, do not necessarily imply that the N400 indexes the use of situational expectancy, as explained above.

This is indeed supported by studies finding that the N400 may be primarily sensitive to association, rather than situational expectancy. For example, in an ERP experiment on Dutch (Hoeks et al., 2004) manipulated world knowledge by contrasting passive sentences as *The javelin was by the athletes thrown* with their implausible active counterparts *The javelin has the athletes thrown*, whose reversed thematic roles violate comprehenders' situational expectations. Interestingly, they do not find a difference in the N400 on *thrown* between those conditions, indicating that the N400 may be mainly reflecting semantic association and not in fact be sensitive to situational expectancy.

Related results come from other studies on these so-called semantic illusions, e.g. Kuperberg et al. (2007), who find no N400 effect on semantically associated, but situationally implausible words in sentences like *At breakfast the eggs would eat*, but rather find a P600 effect (see also Fischler et al., 1983; Kim & Osterhout, 2005; Paczynski & Kuperberg, 2012). The absence of an N400 effect has been explained by plausibility-driven integration processes in multi-stream models (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kuperberg et al., 2007), but may also simply be due to association of the target word with the described event, as pointed out by Brouwer et al. (2012).

Further evidence that association may be sufficient to attenuate the N400 comes from a study by Delogu et al. (2019), who find no N400 effect on *menu* in *John left the restaurant. Before long, he opened the menu …* compared to the same target in *John entered the restaurant. Before long, he opened the menu …*, even though it violates comprehenders' event knowledge. They do, however, find an N400 on the target in *John entered the apartment. Before long, he opened the menu …* compared to their baseline condition. These results can be interpreted as evidence that the N400 can be fully attenuated by the association between *restaurant* and *menu*, regardless of situational expectancy.

Considering these results, the question arises: does the N400 index association or situational expectancy, or is it sensitive to both? To investigate this issue, it is crucial to investigate the relative contribution both these factors may have on N400 amplitude in a single experimental design. This has been attempted in an ERP study by Otten and Van Berkum (2007), who presented participants with event descriptions like *The manager thought that the board of directors should assemble to discuss the issue. He planned a meeting/session where …*, in which the target word was either semantically highly expected (*meeting*) or less expected (*session*), as assessed by a cloze task. Alternatively, the target sentence would follow a slightly altered context version which, for example by negation (*The manager thought that the board of directors need not assemble …*), made the target word an implausible continuation of the event, and hence situationally unexpected. As this also results in a lower cloze value for the formerly high-cloze target *meeting*, we will further refer to the targets as highly associated with the context event (*meeting*) or less associated (*session*) to avoid confusion with the term expectancy. They find an N400 effect for the less associated targets compared to the highly associated targets. However, the scalp distribution of this effect differs depending on whether or not the target was situationally expected or not: when the target was an implausible continuation of the event, the negativity for the less associated target compared to the highly associated target showed a left-lateralised, more central distribution and a lower amplitude, while the negativity for the less associated, but situationally expected targets (compared to the situationally expected and highly associated targets) was broadly distributed and showed a higher amplitude. Otten & Van Berkum interpret their findings as evidence that the N400 reflects situational expectancy as well as semantic expectations, but with possibly different underlying neural generators resulting in different topographic distributions.

These results are interesting, but unfortunately Otten & Van Berkum do not report a direct comparison of the highly associated target (*meeting*) in both plausible and implausible context conditions. This would be especially interesting, as this comparison alone should reveal the influence of situational expectancy: while both context versions are supposed to provide the same semantic expectations for the target word, they differ with respect to the situational expectations. Instead, Otten

& Van Berkum provide results only for the comparison of differences between highly associated and less associated targets in the two context conditions separately. Based on visual inspection, their two highly associated conditions might not differ, but instead the reported effects might be driven almost entirely by the difference in the less associated conditions, again raising the question of how sensitive exactly the N400 is to situational expectancy when association is controlled for. A possible reason for their observation might be that their experimental manipulation of the context, which was supposed to alter comprehenders' situational expectations based on their event knowledge, was not strong enough to produce a reliable effect. A closer look at the type of stimuli used for their study, as provided in their appendix, reveals that only a small number of stimuli used negation to make the context events implausible. Other alternations included more severe changes to the lexical material between the two context versions, for example temporal shifts (*The football players celebrated their victory in the pub./The football players expected to celebrate their victory in the pub later on.*). While a negation marks an event as clearly not happening and makes a continuation of this event implausible, a temporal shift merely marks it as postponed to a (maybe not so distant) future, so a continuation is not necessarily implausible or entirely unexpected. Therefore, in order to make a claim about situational expectancy being reflected in the N400, it is necessary to investigate the effect of event knowledge violations in more thoroughly controlled and homogeneous stimuli.

In two ERP experiments we investigate the influence of both situational expectancy and association on N400 amplitude in a single experimental design. As noted above, expected target words are almost always also associated with the context, and while expectancy can be quantified by measures such as cloze or plausibility judgements, assessing the association of a word with the context is more challenging. In order to dissociate the relative contribution of these two factors, Experiment 1 therefore exploits event scenarios as shown in (2):

(2-a) Kathy played a game of monopoly$_{[\text{active event}]}$. She counted her money and bought Boardwalk$_{[\text{expected target}]}$/cotton candy$_{[\text{unexpected target}]}$.

(2-b) Instead of going to the annual fair$_{[\text{inactive event}]}$, Kathy played a game of monopoly$_{[\text{active event}]}$. She counted her money and bought Boardwalk$_{[\text{expected target}]}$/cotton candy$_{[\text{unexpected target}]}$.

In both scenarios, the critical word is either situationally expected (*boardwalk*) or unexpected (*cotton candy*), given the preceding context sentence: In (2a), the context sentence introduces an ongoing event ("active-only" condition), such that the expected target is also associated with the context, as in many previous studies. In (2b), however, an additional inactive event – explicitly marked as not happening by using the *instead of*-construction[1] – precedes the active event ("active+inactive" condition). Critically, in the unexpected condition, the critical word (*cotton candy*) is one that would be expected if the inactive event were in fact active. As a consequence when the critical word is unexpected, it is still associated with the inactive context event. In the active-only condition however, the inactive event is not mentioned in the context sentence, thus providing no association for the unexpected target word. Importantly, any difference in the degree of expectancy and association of critical words with their contexts is controlled by using a counter-balanced design. That is, for each item as shown in (2), there is a corresponding item in which the active and inactive events – and thus the expected and unexpected target words – are interchanged.

As association may also be influenced by the distance between the event and the target word, Experiment 2 replicates the design for Experiment 1, but presents the active event before the inactive event ("Kathy played a game of monopoly instead of going to the annual fair").

In both experiments, the active-event-only context conditions provide both the upper and lower baseline on the expected N400 effects: we expect the largest N400 effect for the active-event-only/unexpected target condition, where neither situational expectancy nor association provide facilitation of the target word. In contrast, in the active-event-only/expected target condition the target word is both situationally expected and associated with the active event in the context, which is predicted to result in a maximally attenuated N400 amplitude. Unless processing would be hindered by additional lexical material, N400 amplitude should be equally reduced in the active+inactive event/expected target condition, as the target here is also expected and associated with a context event. The active+inactive event/unexpected target condition, however, allows us to dissociate the contribution of association and situational expectancy on N400 amplitude: The influence of situational expectancy should be visible in the amplitude difference between the active +inactive event/unexpected target condition and the active+inactive event/expected target condition:

between these two conditions, association is kept constant by using the same lexical material up to the target word and providing an associated event to either target word in the context. If the N400 indexes situational expectancy, we predict a larger N400 for the situationally unexpected target. If indeed situational expectancy overrides association effects, this N400 effect could be as large as for the active-event-only/unexpected target condition, as both these conditions are equally implausible given the situation. If, however, the N400 mainly indexes association, we would expect no significant difference between the active+inactive event/unexpected target and the active+inactive event/expected target condition, as they provide the same amount of facilitation through association for the target word. Indeed, any amount of facilitation through association that adds to situational expectancy should be visible in the amplitude difference between the active+inactive event/unexpected target and the active-event-only/unexpected target condition, as they are equally implausible, but differ in the amount of association available for the target word.[2]

While our study focuses primarily on N400 effects, former studies have found effects in the P600 time window as well: Metusalem et al. (2012) find a posterior positivity for event-unrelated items that they do not analyse further. Hoeks et al. (2004) find positivities for all anomalous conditions, which they explain by the effort necessary to arrive at a plausible interpretation. Likewise, Delogu et al. (2019) report a P600 for their event-related violation condition compared to their baseline condition, indicating that situational expectancy may be reflected in the P600, but not the N400. In light of these findings we might expect a late positivity for our unexpected target conditions. We will return to this point in the general discussion.

The design for our experiments differs from previous studies in three important aspects: First, we do not use semantic anomalies such as violation of selectional restrictions, as they were used by e.g. Hoeks et al. (2004) and Metusalem et al. (2012). All sentences are perfectly acceptable, except in that they may not fit the situation established by the discourse context sentence. This way we avoid possible confounds that may be visible in the critical time windows. Second, we mark the inactive event unambiguously as not happening by employing an *instead of*-construction and hence its associated – but unexpected – target word as an implausible continuation in every case, reducing variability and uncertainty over the set of stimuli as far as possible, which may have been an issue in previous studies (see Otten & Van Berkum, 2007). Third, our conditions differ based on whether the context provides facilitation of the target word through both situational expectancy and association (expected target conditions), through association only (active + inactive event/unexpected target condition), or no facilitation (active event only/unexpected target condition). This way we are able to dissociate and ultimately quantify the relative contribution both factors may have on the N400 within a single and fully counterbalanced experimental design, which to our knowledge has not been done before.

Table 1. Example of the materials used in Exp 1; the target word is underlined for illustrative purposes only.

Context sentence:	
Introduction	Draußen regnete es. *It was raining outside.*
Active event only	Kathy spielte eine Partie Monopoly. *Kathy played a game of monopoly.*
Active + inactive event	Anstatt zum Jahrmarkt zu gehen, spielte Kathy eine Partie Monopoly. *Instead of going to the annual fair, Kathy played a game of monopoly.*
Target sentence:	
Expected	Sie zählte ihr Geld und kaufte die Schlossallee mit gierigem Blick. *She counted her money and bought Boardwalk with a greedy gaze.*
Unexpected	Sie zählte ihr Geld und kaufte die Zuckerwatte mit gierigem Blick. *She counted her money and bought the cotton candy with a greedy gaze.*

2. Experiment 1

In our first experiment participants read three-sentence stories in which the critical word in the final sentence was either situationally expected or unexpected, given the context, as shown in Table 1. Additionally, the amount of association between the target word and the context varied depending on whether or not the context sentence provided the inactive event that would match the unexpected target. If situational expectancy is reflected in the N400, we predict a difference between the unexpected target and the expected target when both the active and inactive event are present in the context, as the lexical material up to the target word and hence the amount of association is identical across these conditions. If the N400 mainly reflects association, these conditions should not differ significantly. Indeed, if discourse processing is mainly driven by situational expectancy, we do not expect a significant difference between the two unexpected target conditions, irrespective of whether or not the inactive event is mentioned in the context, as they are both equally implausible event continuations.

Only if semantic association is a driving factor also in discourse processing, should we see a reduced N400 on the unexpected target when the semantically associated inactive event is present in the context.

2.1. Method

2.1.1. Participants

Thirty-eight native speakers of German (25 female; mean age: 25, range: 19–33 years), all students at Saarland University, participated in the experiment after giving written informed consent. They were financially compensated for their participation. All were right-handed as assessed by an adapted version of the Edinburgh handedness inventory (Oldfield, 1971) and had normal or corrected-to-normal vision. Six participants had to be excluded from the analysis due to technical errors or excessive artifacts affecting more than 25% of the data.

2.1.2. Materials

We created 70 event scenarios similar to the example shown in Table 1. To avoid any bias towards one of the events in a scenario, each scenario was used a second time, but with the formerly active event serving as the inactive event and vice versa, thus resulting in a total of 140 event scenarios with 4 conditions each, a total of 560 stimuli. Each scenario consisted of three sentences. The first sentence served as an introduction to make the scenario sound more natural.[3] The second sentence contained either one event or two events, with one of them marked as not happening, hence *inactive*, by an *instead of*-construction. The third sentence described a continuation of either the active or the inactive event. The target sentence between conditions only differed on the critical region, the direct object of the sentence.

Following the definition by Schank and Abelson (1977), we defined an event as having a clear temporal boundary and several clearly distinguishable and describable subactions. The two events used in the same scenario were maximally different, such that the target word continuing the inactive event was an implausible continuation of the active event and vice versa, which was validated in an acceptability judgement task (see below). The order of events (inactive event first, active event second) in the context sentence was kept constant over all stimuli. The target sentences always followed the same structure: an introduction of 3–4 words, followed by the main verb, followed by the direct object associated with either the active or the inactive event, and a prepositional phrase to account for sentence wrap-up effects.

Event scenarios were pre-tested in an acceptability judgement task. In an online questionnaire, participants had to read the scenarios and rate their plausibility on a 7-point Likert scale (1 = implausible/very bad, 7 = plausible/very good). All 560 Stimuli were distributed over 8 lists following a latin-square design and interspersed with 120 neutral filler items. 80 German native speakers were recruited via Prolific (www.prolific.com) and paid for their participation. Each list was assigned 10 participants, resulting in 10 ratings per item.

Results are presented in Table 2. On a scale from 1 (very bad) to 7 (very good), items in which the target object was expected were on average rated as plausible with a mean rating of 5.5 (active+inactive event condition) and 5.8 (active-only event condition). Items with an unexpected target object were on average rated as implausible with a mean rating of 2.4 for the active+inactive as well as the active-only event condition. These results indicate that our intended experimental manipulation provides the desired effect: participants rated sentences with an unexpected target as implausible, correctly identifying an event as inactive and not relevant for the situation if it was marked as not happening by an *instead of* construction.

Ten event scenarios with the least difference in ratings between the expected and unexpected targets were discarded from the materials for the EEG experiment along with their corresponding active-inactive event scrambled event scenarios. The remaining 120 sets (480 stimuli total), all listed in the Supplementary Materials, were assigned to 4 lists following a latin-square design, so that each list contained 120 critical items. Each participant saw each event scenario in two conditions such that they were maximally different and no target object was presented twice. Each list was interspersed with the 120 neutral filler items used in the pre-test.

Table 2. Averages (Mean (SD reported in brackets), Median, Mode) of acceptability ratings (7-point scale: 1 = very bad, 7 = very good) from the pre-test for Experiment 1.

Condition		Mean	Median	Mode
Act+inact	Expected	5.5 (1.6)	6	7
Act+inact	Unexpected	2.3 (1.7)	2	1
Act-only	Expected	5.8 (1.5)	6	7
Act-only	Unexpected	2.6 (1.8)	2	1

2.1.3. Procedure

Participants were seated in a dimly lit soundproof booth in front of a 24 inch computer screen. Each trial began with a screen asking the participant for a button press in order to continue. It was followed by a screen displaying the context sentences (introduction and either active-only or active+inactive event condition) as a

whole. In order to continue to the target sentence, participants again had to press a button. Target sentences were presented word-by-word in the centre of the screen, starting with the presentation of a fixation point displayed for 750 ms and followed by the words of the sentence, each presented for 400 ms and preceded by an inter-stimulus-interval, ISI, of 100 ms. Each trial ended with the presentation of a question screen prompting the participant to judge how well the target sentence fitted the context on a 4-point scale. After each judgement there was an ITI of 1000 ms before the next trial began. The experiment was divided into six blocks with a short break after each block. Before the experiment started participants performed a training session consisting of 10 sentences to familiarise them with the task. Experiment sessions lasted approximately 1 hour in total.

2.1.4. Electrophysiological recording

The EEG was recorded from 26 active electrodes, placed according to the international 10–20 system in an elastic cap, using the actiCAP system (Brain Products, Gilching, Germany). AFz served as ground and FCz as reference. The EOG was measured by electrodes placed at the outer canthi of both eyes (horizontal EOG) and above and below the left eye (vertical EOG). Impedances were kept below 5 kOhm for all electrodes. EEG and EOG were amplified using a BrainAmps DC amplifier (Brain Products, Gilching, Germany) and a sampling rate of 500 Hz. No filters were used during the recording.

2.2. Data analysis

The EEG processing was done in Brain Vision Analyzer 2 (Brain Products, Gilching, Germany). The signal was filtered at 0.01–30 Hz and re-referenced to linked mastoids. ERPs were obtained by time-locking each participant's EEG to the target noun and averaging over segments 200 ms pre-stimulus onset to 1200 ms post-stimulus onset with a baseline correction in the 200 ms pre-stimulus onset interval. Trials contaminated by ocular or muscular artifacts (approximately 15%) were discarded before averaging.

Table 3. Averages (Mean (SD reported in brackets), Median, Mode) of acceptability ratings (4-point scale: 1=very good, 2=rather good, 3=rather bad, 4=very bad) for Experiment 1.

Condition		Mean	Median	Mode
Act+inact	Expected	1.4 (0.7)	1	1
Act+inact	Unexpected	3.4 (0.9)	4	4
Act-only	Expected	1.4 (0.6)	1	1
Act-only	Unexpected	3.5 (0.8)	4	4

2.3. Results and discussion

2.3.1. Behavioural data

Average ratings for the behavioural task are given in Table 3.

Items in which the target word was expected where generally rated as "good" or "rather good" (rating of 1 or 2), items in which the target word was unexpected where generally rated as "rather bad" or "bad" (rating of 3 or 4). Within expected and unexpected conditions, there were close to no differences between the active +inactive event condition and the active-event-only condition. The behavioural data thus matches the results from the pre-test and indicates that participants fully processed and understood the scenarios, as they correctly judged the situationally unexpected conditions as unacceptable. Importantly, the data reveals only an influence of situational expectancy, while association does not seem to influence the scenarios' acceptability, as both unexpected conditions are rated as equally bad, despite the act+inact/unexpected conditions still providing an associated event in the context sentence.

2.3.2. N400 time window (300–500 ms)

Figure 1 shows the grand average ERPs on the target word in all conditions for a subset of electrodes. For a more detailed view of the effects on one exemplary electrode (Pz), see Figure 2(a).

As visual inspection did not reveal any visible effects in the P600 time range (600–1000 ms), statistical analysis was restricted to investigate effects in the N400 time window. Statistical analysis was done using mixed-effects models (Baayen et al., 2008) in R (R Core Team, 2018), version 3.3.3, with R-package lme4 (Bates et al., 2015), version 1.1-18-1, with crossed, random effects for participants and items. Following the approach recommended by Barr et al. (2013), we used the maximal model structure containing random intercept and slopes for participants and items. The dependent variable was mean value per single trial in the 300–500 ms time window. To answer the research question which aimed at differences between individual conditions, rather than main effects and interactions, we followed the approach proposed by Schad et al. (2020) in using three planned comparisons as our independent variables. These theoretically motivated comparisons test: (a) the difference between the unexpected vs. the expected target when both the active and inactive event were present in the context sentence; (b) the difference between the unexpected target in the active-only vs. the active+inactive event condition; (c) the difference between the expected target in the active-only vs. the active+inactive event condition;

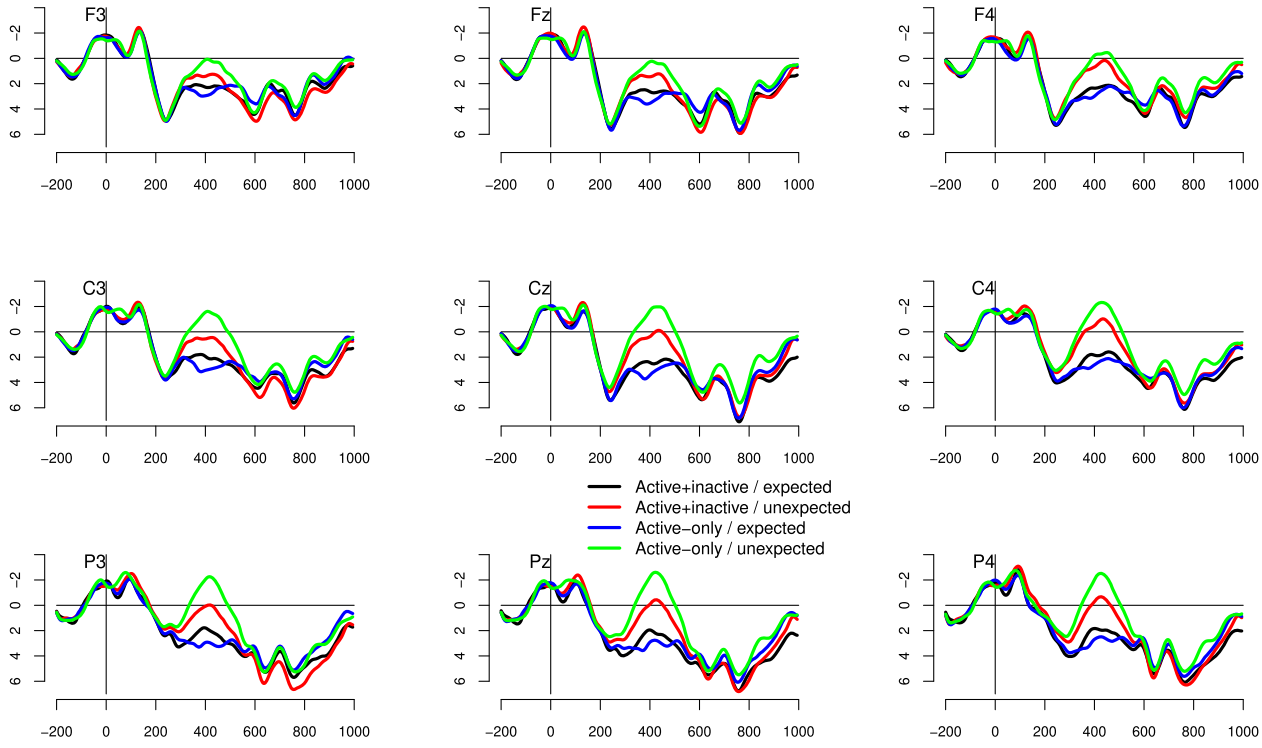
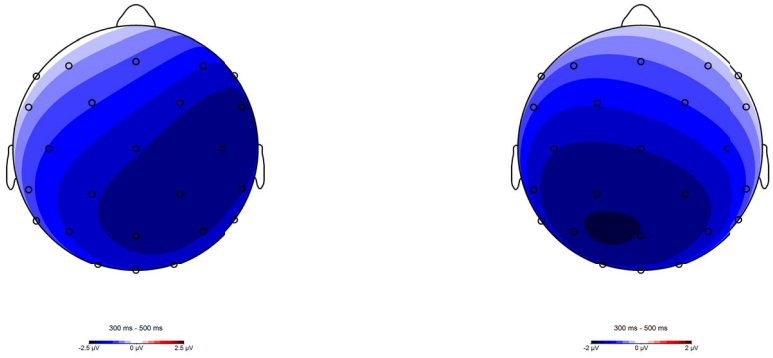


Figure 1. Grand average ERPs time-locked to target noun onset, in experimental conditions active+inactive/expected (black), active +inactive/unexpected (red), active-only/expected (blue), active-only/unexpected (green), in Experiment 1 for a subset of electrodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) The data is filtered at 20 Hz for presentation purposes only, negative voltages are plotted upwards.

This way, comparison (a) tests the effect of situational expectancy when association is controlled for. Comparison (b) tests the effect of association when situational expectancy is controlled for. Lastly, comparison (c) tests whether additional, irrelevant information may interfere with processing facility by comparing the two conditions we believe to be the equally facilitated baseline conditions.

Based on the results from Otten and Van Berkum (2007) there is a possibility that the effects for association and situational expectancy may differ in their topographic distribution. We thus grouped the data from electrode sites into two lateral ROIs, left (F7, F3, FC5, FC1, C3, CP5, CP1, P7, P3, O1) and right (F4, F8, FC2, FC6, C4, CP2, CP6, P4, P8, O2), as well as two longitudinal ROIs, anterior (F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6) and



(a) unexpected vs. expected target in active+inactive context

(b) unexpected target in act-only vs. act+inact context

Figure 2. Topographic maps for Experiment 1 showing the distribution of effects in the N400 time window (300–500 ms). The left panel shows the difference between the active+inactive/unexpected and the active+inactive/expected condition, the right panel shows the difference between the active-only/unexpected and the active+inactive/unexpected condition.

posterior (C3, Cz, C4, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, O2), using deviation coding (−1 vs. 1) for both factors. The model was computed with restricted maximum likelihood estimation.

Model results, including estimates (β), standard error and t-values are provided in Table 4. As an estimation of p-values from mixed models is not straightforwardly possible, we follow Baayen et al. (2008) in considering an absolute *t*-value >2 indicating significance at the 5% level.

First, we find a significant effect for the comparison testing the unexpected vs. the expected target word in the active+inactive event context ($\beta = -1.44$, t=−3.35), showing a larger N400 for the unexpected target word (red line vs. black line in Figure 1 and 2 (a)). The comparison testing the difference between the active+inactive and the active-only event context when the target word is unexpected (green line vs. red line) is significant as well ($\beta = -0.93$, t=−2.69), indicating that the N400 of the unexpected target is attenuated by the inactive, but associated context event. Both ROI factors are significant as well (lateral: $\beta = 0.06$, t=2.11; longitudinal: $\beta = 0.13$, t=4.89). Further, the model reveals a significant interaction of the comparison testing the unexpected vs. the expected target word in the active+inactive event context with lateral ROI ($\beta = 0.42$, t=5.68). This indicates that the negativity for the unexpected vs. the expected target word in the active+inactive event context is more pronounced over the right hemisphere. This distribution of effects is also visible in the topographic maps shown in Figure 3(a). The interaction between the comparison testing the unexpected target in the active-only vs. the active+inactive event condition and lateral ROI is significant as well ($\beta = -0.17$, t=−2.21), revealing that this effect may be larger over the left hemisphere. The interaction with longitudinal ROI is significant for all three comparisons ($\beta = 0.22$, t=2.96; $\beta = 0.32$, t=4.25; $\beta = 0.15$, t=2.03), showing that effects are generally larger over centro-posterior electrode sites than over anterior sites, as is typical for N400 effects. Lastly, the three-way interaction testing the unexpected vs. the expected target in the active+inactive event context × lateral ROI × longitudinal ROI is also significant ($\beta = 0.15$, t=2.04), which indicates that the lateral difference between these two experimental conditions is more pronounced over centro-posterior than over anterior electrode sites. The comparison testing the difference between the expected target word in the active-only vs. the active-inactive event context did not reach significance ($\beta = -0.46$, t=−1.34), thus we cannot assume that interference through additional context information affects the results. Neither the interaction of this comparison with lateral ROI ($\beta = -0.07$, t=−0.98), nor the three-way interaction with lateral and longitudinal ROI ($\beta = 0.02$, t=0.27), nor the Act-only/unexpected vs. act+inact/unexpected × lateral ROI × longitudinal ROI three-way interaction ($\beta = -0.01$, t=−0.06) reached significance.

Table 4. Mixed-effect model results for Experiment 1 in the N400 time window, significant effects printed in bold type.

Effect	Est.	SE	*t*
Act+inact/unexpected vs. act+inact/expected	**−1.44**	**0.43**	**−3.35**
Act-only/unexpected vs. act+inact/unexpected	**−0.93**	**0.35**	**−2.69**
Act-only/expected vs. act+inact/expected	−0.46	0.35	−1.34
ROI (lateral)	**0.06**	**0.03**	**2.11**
ROI (longitudinal)	**0.13**	**0.03**	**4.89**
Act+inact/unexpected vs. act+inact/expected × ROI (lat)	**0.42**	**0.07**	**5.68**
Act-only/unexpected vs. act+inact/unexpected × ROI (lat)	**−0.17**	**0.07**	**−2.21**
Act-only/expected vs. act+inact/expected × ROI (lat)	−0.07	0.07	−0.98
Act+inact/unexpected vs. act+inact/expected × ROI (sag)	**0.22**	**0.07**	**2.96**
Act-only/unexpected vs. act+inact/unexpected × ROI (sag)	**0.32**	**0.07**	**4.25**
Act-only/expected vs. act+inact/expected × ROI (sag)	**0.15**	**0.07**	**2.03**
ROI (lat) × ROI (sag)	0.02	0.03	0.57
Act+inact/unexpected vs. act+inact/expected × ROI (lat) × ROI (sag)	**0.15**	**0.07**	**2.04**
Act-only/unexpected vs. act+inact/unexpected × ROI (lat) × ROI (sag)	−0.01	0.07	−0.06
Act-only/expected vs. act+inact/expected × ROI (lat) × ROI (sag)	0.02	0.07	0.27

2.3.3. Discussion

These results reveal, first, facilitated processing of the expected target word matching the active context event, compared to the unexpected target word matching the inactive context event, as reflected by a significantly reduced N400 for both expected target conditions. This indicates that the N400 indeed reflects situational expectancy. Second, processing of the unexpected target word is facilitated when it is associated with an event presented in the context (active+inactive/unexpected condition), compared to when it is unassociated with the context event (active-only/unexpected condition), indicated by a large N400 for the active-only/unexpected condition. This suggests that even with situational expectancy taken into account, association still contributes to determining N400 amplitude. Interestingly, this is where the ERP results differ from the behavioural results: the acceptability ratings from the ERP experiment and the pre-test show no influence of association. Both unexpected target conditions are clearly rated as unacceptable, despite one of them providing a matching, even if inactive, event in the context. This shows that participants understood

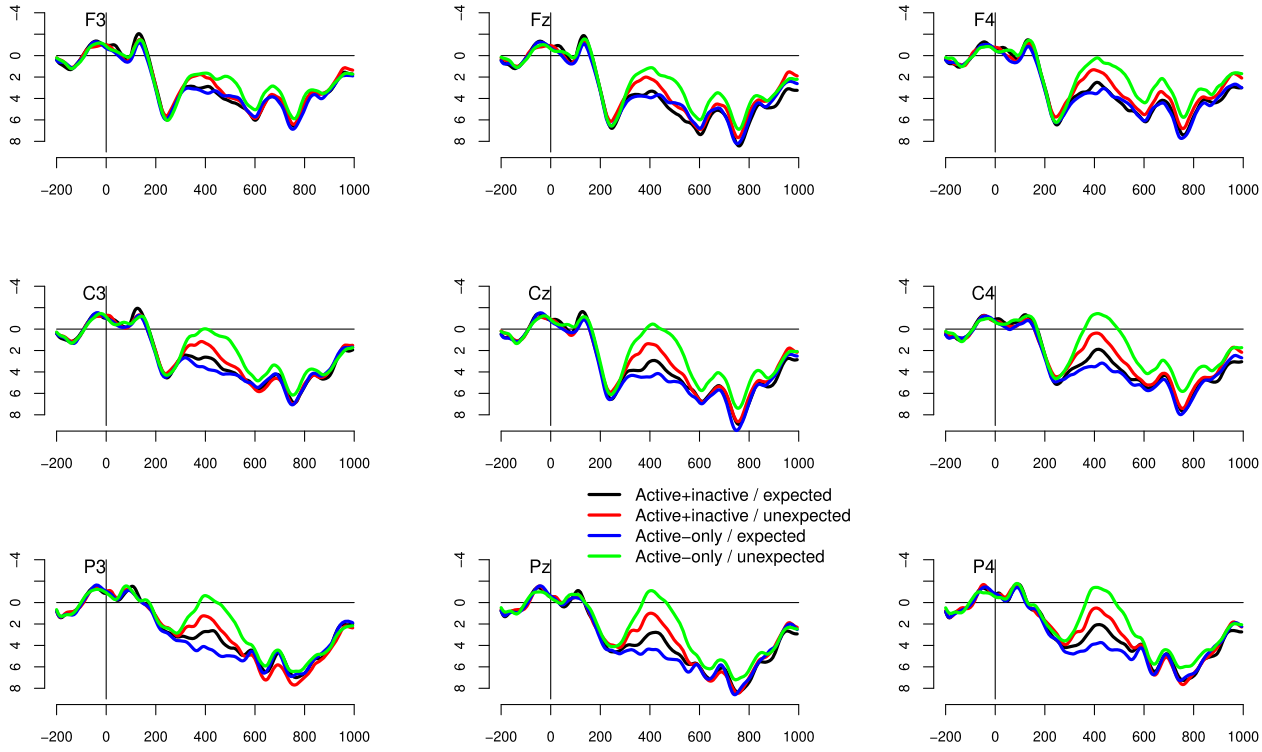


Figure 3. Grand average ERPs time-locked to target noun onset, in experimental conditions active+inactive/expected (black), active +inactive/unexpected (red), active-only/expected (blue), active-only/unexpected (green), in Experiment 2 for a subset of electrodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) The data is filtered at 20 Hz for presentation purposes only, negative voltages are plotted upwards.

the scenarios correctly. Hence, the reduced N400 we attribute to association is unlikely to be explained by accounts like shallow or good enough language processing (Ferreira et al., 2002; Ferreira & Patson, 2007). GE language processing especially would predict a substantial amount of comprehension questions answered incorrectly due to superficial sentence interpretations, which our behavioural data does not support. We will return to this point in the General Discussion. Lastly, the results suggest that effects we attribute to situational expectancy and association may be distributed differently over hemispheres, which has been observed by Otten and Van Berkum (2007) as well. We will return to this point in the general discussion. Together, these results support the hypothesis that both situational expectancy and association are used to facilitate processing rapidly, as both these factors influence N400 amplitude. In order to address a possible concern regarding our interpretation of these results, we ran a follow-up experiment with the goal to replicate the findings from Experiment 1.

Table 5. Example of the materials used in Experiment 2; the target word is underlined for illustrative purposes only.

Context sentence:	
Introduction	Draußen regnete es. *It was raining outside.*
Active event only	Kathy spielte eine Partie Monopoly. *Kathy played a game of monopoly.*
Active + inactive event	Kathy spielte eine Partie Monopoly, anstatt zum Jahrmarkt zu gehen. *Kathy played a game of monopoly instead of going to the annual fair.*
Target sentence:	
Expected	Sie zählte ihr Geld und kaufte die Schlossallee mit gierigem Blick. *She counted her money and bought the park lane with a greedy gaze.*
Unexpected	Sie zählte ihr Geld und kaufte die Zuckerwatte mit gierigem Blick. *She counted her money and bought the cotton candy with a greedy gaze.*

3. Experiment 2

While the results from Experiment 1 are consistent with our hypothesis, the experimental design contains a possible confound regarding our hypothesis about situational expectancy being reflected in N400 amplitude. Some priming studies provide evidence that effects of semantic association may be influenced by the distance between the prime and the target (McNamara, 1992; Ratcliff & McKoon, 1988). In our study, the order of events in the context was kept the same over all stimuli, with the inactive event always being the first one mentioned and the active event always being the last one mentioned. Hence, the distance between the

active event and the target word was always shorter than the distance between the inactive event and the target word. Thus the results we argued were due to situational expectancy might be explainable entirely by semantic association: the large N400 we observed for the active+inactive/unexpected condition (the red line, see Figures 1 and 2(a)) compared to the reduced N400 for the active+inactive/expected condition (black line) might be accounted for by the larger distance between the inactive event and its matching target, and thus, a weaker semantic association effect.

In this follow-up experiment, we adjusted the order of events in the context sentence while keeping all other parameters from Experiment 1. Thus, with the exception of the active event being always mentioned before the inactive event (see Table 5), Experiment 2 is a replication of Experiment 1.

There are two possible outcomes: if our initial hypothesis is correct and both situational expectancy and semantic association contribute to processing facilitation, and both are reflected in N400 amplitude, results between Experiment 1 and Experiment 2 should not differ significantly. If, however, the results from Experiment 1 are in fact due to the recency of the active event, results for Experiment 2 should differ with respect to the active+inactive event conditions: if the N400 effect for the unexpected target condition compared to the expected target condition is explainable by the large distance between the inactive event and its target compared to the short distance between the active event and its target, this effect should vanish or even be reversed for Experiment 2, where the distance between the inactive event and its target is short and the distance between the active event and its target is large.

3.1. Method

3.1.1. Participants

Thirty-seven native German speakers (29 female; mean age: 25, range: 19–31 years), all students at Saarland University, took part in the study after giving written informed consent. None of them had participated in Experiment 1. All participants were right-handed, had normal or corrected-to-normal vision and were paid for their participation. Five participants had to be excluded from the analysis due to excessive artifacts affecting more than 25% of the data. Thirty-two subjects, the same number as was used for the analysis for Experiment 1, remained for the analysis.

Table 6. Averages (Mean (SD reported in brackets), Median, Mode) of acceptability ratings (4-point scale: 1=very good, 2=rather good, 3=rather bad, 4=very bad) for Experiment 2.

Condition		Mean	Median	Mode
Act+inact	expected	1.4 (0.7)	1	1
Act+inact	Unexpected	3.5 (0.9)	4	4
Act-only	Expected	1.3 (0.6)	1	1
Act-only	Unexpected	3.5 (0.8)	4	4

3.1.2. Materials

Experiment 2 used the same lexical material as Experiment 1. Only the order of events used in the active +inactive event conditions was reversed for this study, resulting in the active event always being the first one mentioned in the context sentence and the inactive event being the last one mentioned. So, for example, the context sentence from exp 1 *Instead of going to the annual fair, Kathy played a game of monopoly* was changed to *Kathy played a game of monopoly instead of going to the annual fair* for exp 2. Text passages and fillers, also the same as in Experiment 1, were distributed over 4 list in the exact same way as in Experiment 1.

3.1.3. Procedure and electrophysiological recording

To make the two experiments comparable, the experiment procedure and the technical devices used were identical to Experiment 1, with participants conducting the experiment in a soundproof booth in front of a 24 inch computer screen, following the procedure described in the Methods section for Experiment 1. The EEG was recorded using the system and parameters described for Experiment 1.

3.2. Analyses

The EEG signal was processed using the same parameters as described in the Data Analysis section for Experiment 1. Approximately 14% of data was rejected before averaging.

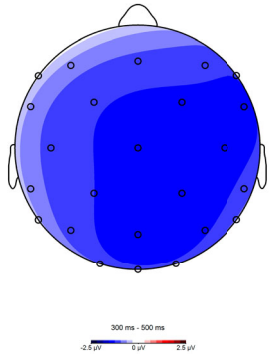
3.3. Results and discussion

3.3.1. Behavioural data

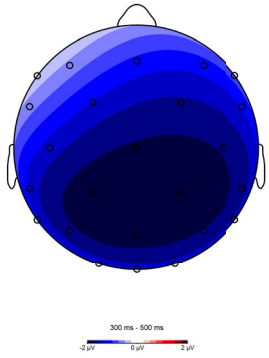
Mean ratings for the behavioural task are given in Table 6. Results did not vary from Experiment 1, with items in which the target word was expected receiving a good rating on average and items in which the target word was unexpected on average being rated as bad. As in Experiment 1, there was no visible influence of association.

3.3.2. N400 time window (300–500 ms)

Figure 4 shows the grand average ERPs on the target word in all conditions on a subset of electrodes.



(a) unexpected vs. expected target in active+inactive context



(b) unexpected target in act-only vs. act+inact context

Figure 4. Topographic maps for Experiment 2 showing the distribution of effects in the N400 time window (300–500 ms). The left panel shows the difference between the active+inactive/unexpected and the active+inactive/expected condition, the right panel shows the difference between the active-only/unexpected and the active+inactive/unexpected condition.

A more detailed view of the effects is provided for one exemplary electrode (Pz) in Figure 2(b).

To make the results for Experiment 1 and Experiment 2 comparable, the same model parameters were used for Experiment 2 as described in the results section for Experiment 1. Due to the similar experiment design, factors, factor levels and factor coding were kept the same. Model results, including estimates (β), standard error, *t*-value and *t*-test results are provided in Table 7.

The results largely replicate the findings from Experiment 1, for a side-by-side comparison of the waveforms see Figure 2. Again, we find a significant effect for the comparison between the unexpected vs. the expected target word in the active+inactive event context ($\beta = -0.98$, t=−2.07), indicating a significantly larger N400 for the unexpected target word (red line vs. black line in Figures 4 and 2(b)). The comparison testing the difference between the active+inactive and the active-only event context when the target word is unexpected (green line vs. red line) is also significant ($\beta = -1.31$, t=−3.13), as reflected in a larger N400 when the inactive event was not mentioned in the context. Both factors for ROI are significant as well (lateral: $\beta = 0.11$, t=3.98; longitudinal: $\beta = 0.13$, t=4.7). Of the interactions with lateral ROI, only the comparison between the unexpected vs. the expected target word in the active+inactive event context was significant ($\beta = 0.2$, t=2.54), again showing that this effect may be more pronounced over the right hemisphere. The interactions between longitudinal ROI and the comparisons testing the unexpected target in both event contexts ($\beta = 0.29$, t=3.73), as well as the comparison testing the expected target in both event contexts ($\beta = 0.4$, t=5.06) are significant, and, as in Experiment 1, indicate that effects may be larger over posterior than over anterior electrode sites.

Different to Experiment 1, however, the interaction of the comparison between the unexpected and the expected target in the active+inactive event context with longitudinal ROI was not significant ($\beta = 0.03$, t=0.38), neither was any of the three-way interactions.

Table 7. Mixed-effect model results for Experiment 2 in the N400 time window, significant effects printed in bold type.

Effect	Est.	SE	*t*
Act+inact/unexpected vs. act+inact/expected	**−0.98**	**0.5**	**−2.07**
act-only/unexpected vs. act+inact/unexpected	**−1.31**	**0.42**	**−3.13**
act-only/expected vs. act+inact/expected	−0.56	0.51	−1.12
ROI (lateral)	**0.11**	**0.03**	**3.98**
ROI (longitudinal)	**0.13**	**0.03**	**4.7**
Act+inact/unexpected vs. act+inact/expected × ROI (lat)	**0.2**	**0.08**	**2.54**
Act-only/unexpected vs. act+inact/unexpected × ROI (lat)	0.14	0.08	1.84
Act-only/expected vs. act+inact/expected × ROI (lat)	0.12	0.08	1.53
Act+inact/unexpected vs. act+inact/expected × ROI (sag)	0.03	0.08	0.38
Act-only/unexpected vs. act+inact/unexpected × ROI (sag)	**0.29**	**0.08**	**3.73**
Act-only/expected vs. act+inact/expected × ROI (sag)	**0.4**	**0.08**	**5.06**
ROI (lat) × ROI (sag)	−0.01	0.03	−0.49
Act+inact/unexpected vs. act+inact/expected × ROI (lat) × ROI (sag)	0.06	0.08	0.8
Act-only/unexpected vs. act+inact/unexpected × ROI (lat) × ROI (sag)	0.08	0.08	0.96
Act-only/expected vs. act+inact/expected × ROI (lat) × ROI (sag)	0.06	0.08	0.75

3.3.3. Discussion

The results from Experiment 2 largely replicate the central findings from Experiment 1. Again we find a graded N400 effect, with a reduced N400 effect for expected target words, irrespective of whether or not the context provided an additional inactive event. We find an intermediate N400 for the unexpected target word when the matching inactive event was presented

in the context, and the largest N400 effect for the unexpected target word when only the active event was presented in the context. If the N400 effect observed for the active+inactive/unexpected condition in Experiment 1 was explainable by weaker semantic association due to a recency effect, this comparison should not have been significant in the statistical analysis for Experiment 2. However, we still find this comparison to be significant in Experiment 2, despite the shorter distance between the inactive context event and the unexpected target. We thus confirm our hypothesis that both situational expectancy and association are reflected in the N400 and the results from our first experiment are not explainable by either expectancy or association alone.

4. General discussion

While evidence in the literature has shown that ultimately, event knowledge is crucial for language comprehension, ERP studies investigating the on-line use of event knowledge are often unable to distinguish between aspects of association and situational expectancy. Association between two words may be formed by our knowledge about events, and thus participants and objects likely to appear within them. Situational expectancy denotes precise expectations about the unfolding situation described in the discourse. Hence it remained unclear how sensitive the N400 is to situational expectancy, above and beyond simple association. In two ERP experiments we investigated the relative influence both situational expectancy and association have on N400 amplitude in a single experimental design, which to our knowledge has not been done before.

Following a context sentence with either one event marked as active/ongoing, or an additional event marked as inactive/not happening, we presented a target sentence in which the target word continued either the active or the inactive event. Thus, the target word was either situationally expected or unexpected, depending on whether it continued the ongoing event or the event clearly marked as not happening. Additionally, association between the unexpected target word and the inactive event was manipulated by varying whether the inactive event was mentioned in the context sentence or not: while the active event always provided association for the expected target word, the unexpected target word could only be associated with the inactive event when it was presented.

We find a reduced N400 effect on the expected target words, a larger N400 effect on the unexpected target when the inactive event was mentioned in the context (thus controlling for association between these conditions), and the largest N400 effect on the unexpected target when only the active event was presented in the context, thus providing no association for the target word in this condition compared to all other conditions. Our results hence suggest that the N400 is sensitive to situational expectancy, as we find an N400 effect for situationally unexpected words compared to expected words, even when association was controlled for. Equally, however, the N400 also reflects association, as the N400 effect was largest when the target word was not only unexpected but also not associated with the context.

In the second experiment we reversed the order of events in the active+inactive event context conditions, thereby manipulating the distance between an event and its matching target word across experiments. This was done to avoid a possible confound in Experiment 1, where the effect we explained by influences of situational expectancy might also be explained by a recency effect, and hence purely by semantic association. The results, replicating the findings from our first experiment, suggest that N400 amplitude indeed robustly reflects situational expectancy as well as association, as the observed effects cannot be attributed to influences of semantic association alone.

Importantly, the graded ERP results do not pattern with the acceptability ratings both from the pre-test of the materials and the task used in the ERP experiments. There, participants rated scenarios with the unexpected target as equally bad, irrespective of whether or not the context sentence provided the matching inactive event or not. Thus, the plausibility of the target word alone does not predict the observed N400 pattern. At the same time the behavioural data strongly suggests that the ERP results cannot be explained by shallow processing: participants clearly understood the scenarios correctly and were aware of the situational implausibility in the unexpected conditions. Good-enough or shallow processing hinges on the idea that sentences may often be misinterpreted, as comprehenders may analyse input only superficially. Thus, an implausible target word may not be recognised as implausible if it is semantically connected to the context, reflected in a suppressed N400 effect. The underspecified processing may stem from prediction mechanisms, where strong prior expectations override analysis of the current input (Kuperberg & Jaeger, 2016), or from information structured into given/new (Ferreira & Lowder, 2016), or the inherently automatic, probabilistic and stimulus-driven nature of the underlying process (Rabovsky et al., 2018). It is tempting to explain the reduced N400 effect for the active+inactive/unexpected condition in our study under such an account. However,

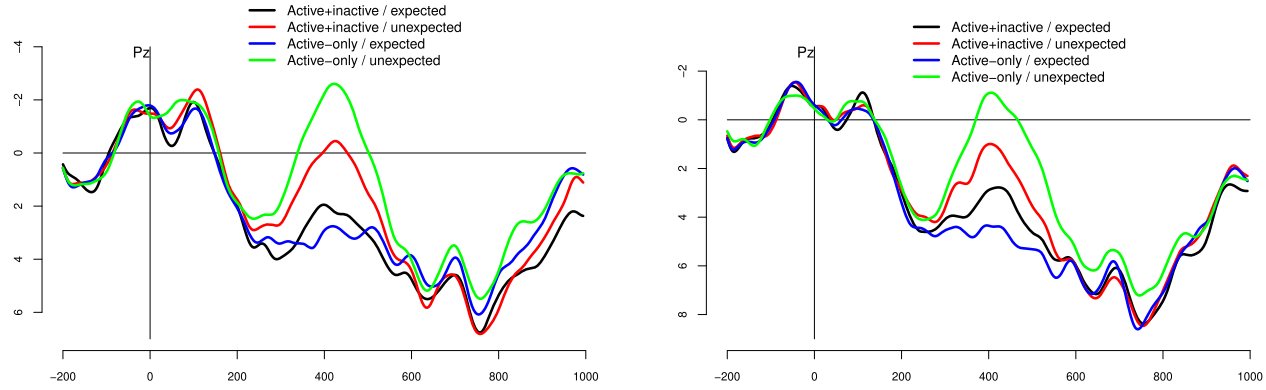
studies reporting effects of shallow or heuristic processing usually involve the existence of strong, semantically attractive target words (such as *For breakfast the eggs would eat …*, Kuperberg et al., 2007) distracting from the correct interpretation, which is not the case in our studies. The implausibility of the target word stems from the *instead of*-construction in the context sentence. There is no good reason to assume that comprehenders would not correctly process and understand this, as it is presented self-paced, giving them enough time for processing. This is reflected in the behavioural data, as participants successfully rejected implausible sentences in the judgement task. Thus, our data does not support GE language processing accounts claiming that the inaccurate interpretation may not be resolved, which should be reflected in inaccurately answered comprehension questions. In order to not detect the implausible target word, comprehenders would then have to somehow re-interpret and misunderstand the context, but then arrive at a correct sentence interpretation again to correctly answer the comprehension question.

Our findings are in line with a great body of evidence from the literature stating that even in initial processing stages, event knowledge may be used to facilitate processing (Bicknell et al., 2010; Camblin et al., 2007; Chwilla & Kolk, 2005; Delogu et al., 2018; Metusalem et al., 2012; Otten & Van Berkum, 2007). However, most of the studies to date have not considered the role of the N400 component and its sensitivity to different factors in language processing that may lead to the observed facilitation effect. Hence, many of the findings in the literature could also have been attributed simply to association between words, possibly related via events, and not necessarily require very detailed mental models of the situation (Chwilla & Kolk, 2005; Metusalem et al., 2012). Even more so, some studies suggest that the N400 may be mainly sensitive to association and not be influenced by the plausibility of the situation (Delogu et al., 2019; Hoeks et al., 2004). In our experiments we specifically dissociated association from situational expectancy and were able to show that both these factors are reflected in the N400 component.

Regarding the topographic distribution of N400 effects to manipulations of association and expectancy, Otten and Van Berkum (2007) find the N400 effect for less associated vs. highly associated words in implausible event continuations to be more left-lateralised and centrally distributed than the broadly distributed N400 effect for less associated vs. highly associated words in plausible event continuations. They thus argue that semantic expectancy (which somewhat refers to what we call association) and discourse expectancy (which is what we refer to as situational expectancy) may be at least partly different processes. In our experiments, the comparison most closely matching their manipulation is the active+inactive event/unexpected target vs. active-only event/unexpected target, which tests for effects of association when the target is situationally unexpected. Based on visual inspection, the effect may indeed be slightly left-lateralised in our first experiment, but evenly distributed over both hemispheres in the second experiment (see Figures 3(b) and 5(b)). In our statistical analysis the interaction of this comparison with ROI is significant in Experiment 1, but not in Experiment 2 (see Tables 4 and 7). As the hypothesis that association effects may be larger when distance between the active event and its target is reduced was the premise for Experiment 2, one can only speculate why this effect is not robust in our second experiment. One possible explanation may be that effects in general appear to be less pronounced in Experiment 2 compared to Experiment 1. As for the topographical distribution of effects we do, however, find a significant interaction between hemisphere and the comparison testing the effect of situational expectancy when association is controlled for in our first experiment (see Table 4 and Figure 3(a)), indicating that the effect of situational expectancy may be more right-lateralised. Interestingly, this finding is in line with the effects reported by Metusalem et al. (2012), even though we argue those may also be attributable to association. Again, this finding does not hold for our second experiment, where this interaction is not significant. As these lateralisation effects do not appear to be particularly robust, and other studies report the typical broadly distributed N400 effects (Bicknell et al., 2010; Delogu et al., 2019; Hoeks et al., 2004), we refrain from speculations on the functional interpretation of left- and right-lateralised N400 effects, although this may be worth further investigation in future studies.

Considering possible effects of interference, neither Experiment 1 nor Experiment 2 revealed a significant effect of the inactive event being present in the context when the target word was expected. This suggests, within the limits of our design, that the amount of context per se does not play a significant role in processing difficulty. As long as the target word matches the context event that has been marked as active or ongoing, mentioning additional, inactive context events does not seem to interfere with processing difficulty.

With respect to late positive effects, P600 effects in response to semantic violations have increasingly been reported in the literature (for an overview, see Van Petten & Luka, 2012). and also in studies investigating event knowledge. Specifically, Hoeks et al. (2004) find a



(a) Experiment 1 (b) Experiment 2

Figure 5. Grand average ERPs time-locked to target noun onset, in experimental conditions active+inactive/expected (black), active +inactive/unexpected (red), active-only/expected (blue), active-only/unexpected (green), for Experiment 1 (left) and Experiment 2 (right) for one exemplary electrode (Pz). The data is filtered at 20 Hz for presentation purposes only, negative voltages are plotted upwards. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

P600 for their anomalous conditions, which they interpret as an indicator for the effort necessary to arrive at a plausible sentence interpretation. Metusalem et al. (2012) find a P600 for both their event-related and their event-unrelated violation condition, but do not comment further on it. Delogu et al. (2019) report a significant P600 for their event-related violation. Thus, in our experiments we may have expected a positivity for the situationally unexpected target words. However, visual inspection of the data did not indicate a reliable P600 effect for any of our conditions. There is evidence that such positivities can often be masked by a preceding N400 due to component overlap (Delogu et al., 2021). We refrain from further speculation, as the presence of a P600 effect is not the focus of this study.

Regarding the findings from Otten and Van Berkum (2007), we specifically targeted some weak points in their study by using well controlled, homogeneous stimuli with the inactive event unambiguously declared as not happening, and hence making its continuation implausible in every single item, while keeping the lexical material within item sets identical up to the target word. This way we hoped to maximise the difference between our plausible and implausible conditions and minimise variability and uncertainty over items. While their design provided an opportunity to directly test the influence of discourse expectations on N400 amplitude when semantic expectation was controlled, they did not report this comparison. We are thus able to expand on their findings by demonstrating that even when association is controlled for, situational expectancy still also modulates N400 amplitude.

We also eliminated a possible confound some studies implemented in their design when using semantic anomalies to investigate influences of world knowledge (Hoeks et al., 2004; Metusalem et al., 2012). For example, the N400 effects on the event-related and event-unrelated targets compared to the baseline in Metusalem et al. (2012) may be explained partially by association, but partially by the selectional restrictions making the target words unsuitable arguments for the verbs used. By using syntactically and semantically well-formed sentences we ensured that our experimental manipulations only differed with respect to the amount of association and situational expectancy available for processing of the target word. Thus, our findings provide clearer evidence for the claim by Metusalem et al. (2012) that detailed event knowledge expectations are reflected in N400 amplitude.

While our study was not designed to address the functional interpretation of the N400 – since expectancy effects are consistent with several accounts, including retrieval, integration, and hybrid accounts – the finding that association is also reflected in N400 amplitude is difficult to reconcile within an integration account: from an integration point of view, an unexpected target should be difficult to integrate into the context regardless of its association with the prior context. Importantly, an influence of situational expectancy on the N400 does not imply that the N400 indexes semantic integration. It merely implies that situational expectancy influences word expectancy, which can be explained by retrieval accounts as well as, by extension, the retrieval mechanism in hybrid accounts.

Since our findings demonstrate that situational expectancy may be reflected in N400 amplitude, the question remains as to why other studies do not find N400 effects of situational expectancy. For example,

Delogu et al. (2019) find an N400 effect on *menu* in *"John entered the apartment. Before long, he opened the* menu …*"* compared to *"John entered the restaurant. Before long, he opened the* menu …*"*, which can be explained by (lack of) association. According to our findings, one could argue that their third condition, *"John left the restaurant. Before long, he opened the* menu …*"*, should evoke an N400 effect as well, as *menu* in this case is associated with the context, but situationally unexpected.

First, we want to point out that we believe the manipulation in our study to be stronger: the inactive event in our study reduces situational expectancy for the target word to a greater extent than the event marked as merely completed in Delogu et al. (2019). An effect of situational expectancy may therefore not be reliably visible in the N400, as we have argued in the case of the study by Otten and Van Berkum (2007) above.

Second, our study does not imply that both association and situational expectancy are equally relevant for processing in any situation. Rather, we propose that they interact dynamically based on whatever proves most helpful in facilitating meaning retrieval. Under this account, situational expectancy may take priority in cases where an elaborate context allows for a detailed model of the situation, but association may take precedence when context information is limited. For example, in our experiments we used a somewhat elaborate context introducing up to two events. In this case, association may provide facilitation for the target word both when it continued the active and when it continued the inactive event. Only situational expectancy, however, could provide further facilitation for the expected target word. In contrast, Delogu et al. (2019) used only a short context sentence introducing one event, while Hoeks et al. (2004) only used single sentences in their design. In both studies, while the target word did not match situational expectations, there was no more expected alternative provided by the context. In those cases, detailed situational knowledge may not be the most reliable source to facilitate retrieval, leading retrieval mechanisms to rely more on association.

5. Conclusion

In conclusion, the present findings provide further insight into the role of the N400 component in language processing. Former studies investigating the role of event knowledge on on-line language processing have often confounded different factors that contribute to event knowledge, i.e. association between words or concepts and situational expectations arising from the context. In two ERP experiments we demonstrated that both association and situational expectancy are simultaneously reflected in N400 amplitude and thus influence word expectancy. This is especially interesting as the behavioural data only shows effects of expectancy. We expand on former studies by providing a more fine-grained look on the contributing factors to N400 effects that so far have been subsumed under the general notion of event knowledge. This has implications for the functional interpretation of the N400 component as well: if the N400 is set up to reflect situational expectancy, which is differently explained under retrieval and hybrid/integration accounts, association needs to be controlled for. Our findings are in line with current accounts in the literature and can, for example, be naturally explained under a lexical retrieval account of the N400. Thus, for future studies of event knowledge, ERP components and the language models they inform, our findings underline the importance to distinguish between the different factors that may independently impact N400 amplitude.

Notes

1. The Penn Discourse Treebank corpus classifies *instead of* as discourse connective of the "chosen alternative" type, stating "the connective indicates that two alternatives are evoked in the discourse but only one is taken" (Prasad et al., 2008, section 4.6.3). Hence, an event marked as not happening this way may be linguistically salient, but situationally rejected.
2. It should be noted that they also differ in the amount of lexical material provided. Although we do not necessarily assume any interference from an unrelated event, we aim for a minimal difference between our conditions. Thus we decided to omit the inactive event, rather than matching the context sentences for length by using an inactive event unrelated to the unexpected target word, in order to not introduce any association-related effects that cannot be fully controlled for.
3. These sentences were designed to be as general as possible, but also consistent with the active event. They may therefore in some cases provide additional association or expectancy for the expected target word, which we have not explicitly controlled for. Implicitly, however, this variation between items is captured in the statistical model within the random effects structure, and is also minimised by our counterbalanced design.

Acknowledgments

We thank Francesca Capuano, Nadja Fürst and Hong Hanh Pham for their assistance with data collection, as well as Phillip M. Alday for helpful comments on the statistical models.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 232722074 – SFB 1102 "Information Density and Linguistic Encoding".

References

Baayen, R., Davidson, D., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baggio, G. (2018). *Meaning in the brain*. MIT Press.

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. https://doi.org/10.1080/01690965.2010.542671

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bicknell, K., Elman, J. L., Hare, M., McRae, K., & Kutas, M. (2010). Effects of event knowledge in processing verbal arguments. *Journal of Memory and Language*, *63*(4), 489–505. https://doi.org/10.1016/j.jml.2010.08.004

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73. https://doi.org/10.1016/j.brainresrev.2008.05.003

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143. https://doi.org/10.1016/j.brainres.2012.01.055

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*(1), 34–44. https://doi.org/10.1162/jocn.1993.5.1.34

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language*, *56*(1), 103–128. https://doi.org/10.1016/j.jml.2006.07.005

Chwilla, D. J., & Kolk, H. H. (2005). Accessing world knowledge: Evidence from N400 and reaction time priming. *Cognitive Brain Research*, *25*(3), 589–606. https://doi.org/10.1016/j.cogbrainres.2005.08.011

Coulson, S., Federmeier, K. D., Van Petten, C., & Kutas, M. (2005). Right hemisphere sensitivity to word-and sentence-level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 129–147. https://doi.org/10.1037/0278-7393.31.1.129

Delogu, F., Brouwer, H., & Crocker, M. W. (2019). Event-related potentials index lexical retrieval (N400) and integration (P600) during language comprehension. *Brain and Cognition*, *135*(1), 103569. https://doi.org/10.1016/j.bandc.2019.05.007

Delogu, F., Brouwer, H., & Crocker, M. W. (2021). When components collide: Spatiotemporal overlap of the N400 and P600 in language comprehension. *Brain Research*, *1766*(9), 147514. https://doi.org/10.1016/j.brainres.2021.147514

Delogu, F., Drenhaus, H., & Crocker, M. W. (2018). On the predictability of event boundaries in discourse: An ERP investigation. *Memory & Cognition*, *46*(2), 315–325. https://doi.org/10.3758/s13421-017-0766-4

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, *33*(4), 547–582. https://doi.org/10.1111/j.1551-6709.2009.01023.x

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15. https://doi.org/10.1111/1467-8721.00158

Ferreira, F., & Lowder, M. W.. (2016). Prediction, information structure, and good-enough language processing. In Brian Ross (Ed.), *Psychology of learning and motivation* (Vol. 65, pp. 217–247). Academic Press.

Ferreira, F., & Patson, N. D. (2007). The "good enough" approach to language comprehension. *Language and Linguistics Compass*, *1*(1–2), 71–83. https://doi.org/10.1111/j.1749-818X.2007.00007.x

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, *44*(4), 516–547. https://doi.org/10.1006/jmla.2000.2728

Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry Jr, N. W. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, *20*(4), 400–409. https://doi.org/10.1111/j.1469-8986.1983.tb00920.x

Golden, R. M., & Rumelhart, D. E. (1993). A parallel distributed processing model of story comprehension and recall. *Discourse Processes*, *16*(3), 203–237. https://doi.org/10.1080/01638539309544839

Hagoort, P. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304* (5669), 438–441. https://doi.org/10.1126/science.1095455

Hare, M., Jones, M., Thomson, C., Kelly, S., & McRae, K. (2009). Activating event knowledge. *Cognition*, *111*(2), 151–167. https://doi.org/10.1016/j.cognition.2009.01.009

Hoeks, J. C., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, *19*(1), 59–73. https://doi.org/10.1016/j.cogbrainres.2003.10.022

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, *5*(4), 281–312. https://doi.org/10.1080/01690969008407065

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*(2), 205–225. https://doi.org/10.1016/j.jml.2004.10.002

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence

from eye tracking. *Cognitive Science*, *30*(3), 481–529. https://doi.org/10.1207/s15516709cog0000_65

Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language*, *57*(4), 519–543. https://doi.org/10.1016/j.jml.2007.01.003

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*(1), 95–127. https://doi.org/10.1016/j.cognition.2004.03.002

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D. N., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*, *100*(3), 223–237. https://doi.org/10.1016/j.bandl.2005.12.006

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163. https://doi.org/10.1038/307161a0

Kutas, M., & Hillyard, S. A. (1989). An electrophysiological probe of incidental semantic association. *Journal of Cognitive Neuroscience*, *1*(1), 38–49. https://doi.org/10.1162/jocn.1989.1.1.38

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933. https://doi.org/10.1038/nrn2532

Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory and Cognition*, *31*(5), 663–672. https://doi.org/10.3758/BF03196106

McNamara, T. P. (1992). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1173–1190. https://doi.org/10.1037/0278-7393.18.6.1173

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. R. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition*, *33*(7), 1174–1184. https://doi.org/10.3758/BF03193221

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*(3), 283–312. https://doi.org/10.1006/jmla.1997.2543

Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & J. L. Elman (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of Memory and Language*, *66*(4), 545–567. https://doi.org/10.1016/j.jml.2012.01.001

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (Vol. 11, pp. 264–336). Lawrence Erlbaum.

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., & Donaldson, D. I. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1791), 20180522. https://doi.org/10.1098/rstb.2018.0522

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*(7), 1098–1111. https://doi.org/10.1162/jocn.2006.18.7.1098

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Otten, M., & Van Berkum, J. J. A. (2007). What makes a discourse constraining? Comparing the effects of discourse message and scenario fit on the discourse-dependent N400 effect. *Brain Research*, *1153*(1), 166–177. https://doi.org/10.1016/j.brainres.2007.03.058

Paczynski, M., & Kuperberg, G. R. (2012). Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *Journal of Memory and Language*, *67*(4), 426–448. https://doi.org/10.1016/j.jml.2012.07.003

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A. K., & Webber, B. L. (2008). The Penn Discourse TreeBank 2.0. In Lrec.

Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*, *2*(9), 693–705. https://doi.org/10.1038/s41562-018-0406-4

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*(3), 385–408. https://doi.org/10.1037/0033-295X.95.3.385

R Core Team. (2018). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria. https://www.R-project.org/

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*(4), 104038. https://doi.org/10.1016/j.jml.2019.104038

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquire into human knowledge structures*. Lawrence Erlbaum.

Swaab, T. Y., Baynes, K., & Knight, R. T. (2002). Separable effects of priming and imageability on word processing: An ERP study. *Cognitive Brain Research*, *15*(1), 99–103. https://doi.org/10.1016/S0926-6410(02)00219-7

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 443–467. https://doi.org/10.1037/0278-7393.31.3.443

Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*(6), 657–671. https://doi.org/10.1162/089892999563724

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes*, *8*(4), 485–531. https://doi.org/10.1080/01690969308407586

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Van Petten, C., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociates lexical and sentential context effects. *Psychological Science*, *8*(3), 238–242. https://doi.org/10.1111/j.1467-9280.1997.tb00418.x

Venhuizen, N. J., Crocker, M. W., & Brouwer, H. (2019). Expectation-based comprehension: Modeling the interaction of world knowledge and linguistic experience. *Discourse Processes*, *56*(3), 229–255. https://doi.org/10.1080/0163853X.2018.1448677