About edible restaurants: conflicts between syntax and semantics as revealed by ERPs

Miriam Kos1*, Theo Vosse1,2, Daniëlle van den Brink1,3 and Peter Hagoort1,4*

1 Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands
2 Department of Psychology, Leiden University, Leiden, Netherlands
3 Behavioral Science Institute, Radboud University Nijmegen, Nijmegen, Netherlands
4 Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

INTRODUCTION

In recent years event related potentials (ERPs) have been used quite extensively to study multiple aspects of language processing. In the past, distinct ERP effects were observed for semantic and syntactic processing. The ERP effects of interest here are the N400 and the P600 effect. Modulations of the N400 amplitude are generally believed to reflect the processing costs of integrating information that is built up on the basis of the preceding language input (Hagoort, 2003). However, others propose that the N400 reflects the ease with which information can be accessed from semantic memory (Kutas and Federmeier, 2000; Federmeier and Laszló, 2009). Words that are semantically incongruent or have a poor semantic fit given the preceding context elicit a larger N400 than words that fit well within the context. This context can either be a single word (Brown and Hagoort, 1993), a sentence (Kutas and Hillyard, 1980, 1984; Hagoort and Brown, 1994), or a discourse (van Berkum et al., 1999b, 2003). The P600 effect (in the past also labeled the syntactic positive shift, SPS) has been ascribed to syntactic processing, since it has been shown to be elicited by syntactic violations (Hagoort et al., 1993) and by syntactically ambiguous structures (Osterhout and Holcomb, 1992; Osterhout et al., 1994; van Berkum et al., 1999a). It has also been shown that this effect is modulated by syntactic complexity (Kaan et al., 2000). It is assumed that the P600 is a reflection of syntactic unification, in which the amplitude of the P600 is affected by competition between alternative unification options (Hagoort, 2003). Others suggest that the P600 reflects syntactic reanalysis or repair (Friederici, 2002).

In order to investigate conflicts between semantics and syntax, we recorded ERPs, while participants read Dutch sentences. Sentences containing conflicts between syntax and semantics (Fred eats in a sandwich…/Fred eats a restaurant…) elicited an N400. These results show that conflicts between syntax and semantics not necessarily lead to P600 effects and are in line with the processing competition account. According to this parallel account the syntactic and semantic processing streams are fully interactive and information from one level can influence the processing at another level. The relative strength of the cues of the processing streams determines which level is affected most strongly by the conflict. The processing competition account maintains the distinction between the N400 as index for semantic processing and the P600 as index for structural processing.

Keywords: language, sentence processing, conflict, semantics, syntax, ERPs

The picture that emerged on the basis of these studies seemed to be quite clear and simple. The P600 effect is related to syntactic processing, whereas the N400 component is a reflection of semantic processing. However, recently there have been several ERP studies which seem to be at variance with the general consensus that P600 and N400 effects are related to syntactic and semantic processes respectively. These studies investigated conflicts in thematic role assignment within the argument structure of the verb. Violations that were seemingly semantic in nature elicited P600 effects instead of N400 effects (Kolk et al., 2003; Kuperberg et al., 2003, 2007; Hoeks et al., 2004; Kim and Osterhout, 2005; Nakano et al., 2010). For example, in the syntactically well-formed sentence “Every morning at breakfast the eggs would eat…” (Kuperberg et al., 2003), it could be expected that the verb “eat” would elicit an N400, as on the basis of syntactic parsing it is inferred that “eggs” is the agent. This is in conflict with the semantic–thematic knowledge that “eggs” do not eat. However, the verb “eat” elicited a P600 effect (Kuperberg et al., 2003). Other examples in which comparable P600 effects were obtained are: “De kat die voor de muizen vluchtte…” (Kolk et al., 2003) and “The cat that fled from the mice…” (Kolk et al., 2003) and “The pizza had been delivering….” (Kim and Osterhout, 2005).

A difference between these sentences and the standard N400 paradigm is that in all these sentences the nouns and the verbs form a plausible combination of content words – for instance morning, breakfast, eggs, and eat (Kuperberg et al., 2003). As a consequence, a plausible script can be formed on the basis of these content words – eggs are eaten at breakfast in the morning –, whereas in the standard N400 sentence processing paradigms such plausible relations are usually lacking (e.g., “He spread the warm bread with socks…”).
(Kutas and Hillyard, 1980; Hagoort et al., 2004). The violation becomes apparent via a conflict with the syntactic structure, which tells us that eggs are the agent of the verb “eat.”

These findings have been incorporated within several language processing models: the Monitoring Theory (Kolk et al., 2003; van Herten et al., 2005, 2006; van de Meeren et al., 2009, 2010), the extended argument dependency model (eADM; Bornkessel and Schlesewsky, 2006; Bornkessel-Schlesewsky and Schlesewsky, 2008, 2009), the continued analysis account (Kuperberg, 2007; Kuperberg et al., 2008), the semantic attraction account (Kim and Osterhout, 2005), and the processing competition account (Hagoort et al., 2009). All these models assume that language processing proceeds along (at least) two processing routes and it is the conflict between these processing streams that leads to the P600 effect in the example sentences displayed above. The way these streams interact differs between the several theories. The Monitoring Theory and the eADM assume that, after an initial phase of independent parallel processing, a conflict between the processing streams subsequently triggers a general cognitive process, which is labeled monitoring (van de Meeren et al., 2009) or generalized mapping (Bornkessel-Schlesewsky and Schlesewsky, 2008) respectively. Kuperberg et al. (2007) assume that the parallel processing streams are fully interactive and she gives a combinatorial interpretation for the P600 effect. Also Kim and Osterhout assume interaction between the parallel processing streams and they maintain a structural interpretation of the P600. They state that the P600 effect is the result of semantic attraction within the sentence (Kim and Osterhout, 2005). For example, when readers encounter the sentence “the pizza had been delivering…”, the semantic attraction to interpreting pizza as the theme of the sentence is so compelling that this dominates the syntactic analysis. Because of this semantic attraction the language comprehender perceives a syntactically well-formed string to be syntactically ill-formed. This results in a P600 effect (Kim and Osterhout, 2005). In the same line of reasoning Hagoort and colleagues assume that semantics, in general, can influence syntactic processing, within a language system consisting of a parallel architecture (Jackendoff, 2002). This system is highly interactive in the sense that the constraints of the independent systems of syntax and semantics are taken into consideration concurrently during listening or reading (Hagoort, 2003, 2005). Information from one level can influence the processing at another level. A conflict between semantic and syntactic constraints is signaled via an interface. The system works according to the principle that relatively stronger cues from one level can drive the processing costs at the other level (Hagoort et al., 2009) and early delivered cues take precedence over incrementally later delivered cues. Hence, the relative strength of the cues determines which level gets the extra processing costs. In case of a conflict, the processing level with the strongest cues imposes an extra processing load at the processing level with the weaker cues, for instance in the case of “The pizza had been delivering” (Kim and Osterhout, 2005) the strong semantic relation between pizza and deliver leads to a higher processing load at the syntactic stream. We will refer to this explanation as the processing competition account.

The examples mentioned above (e.g., The pizza had been delivering; Kim and Osterhout, 2005) all contain strong semantic cues. According to the processing competition account these stronger semantic cues lead to a higher processing load at the syntactic level and thus elicit a P600 effect (Hagoort et al., 2009), which reflects the effortful process of assigning grammatical roles (Wassenaar and Hagoort, 2007). The processing competition account also predicts, in cases where strong syntactic and weak semantic cues are combined, that conflicts between structural and plausible role assignments become apparent at the semantic level, thus eliciting an N400 effect. Importantly, this account maintains the distinction between the N400, as an index of semantic processing, and the P600 as an index of structural processing.

In this paper we investigated sentences, which contain plausible scripts, but are in conflict with the syntactic structure, to see whether these conflicts elicit N400 effects, instead of a P600 effect as was found in previous experiments (Kolk et al., 2003; Kuperberg et al., 2003; Kim and Osterhout, 2005). An example is shown in Table 1.

In these sentences we used transitive verbs for which the presence of a direct object is optional. Two different sentence structures served as a baseline [e.g., (1) Fred eats a sandwich…, (2) Fred eats a restaurant…]. In the first example sentence (NP–V–NP) the verb takes a direct object. This argument consists of a noun phrase which typically is a patient in relation to the verb. Hence, “sandwich” is usually the patient of the verb “to eat.” In the second example sentence the verb is followed by an adjunct. Again we chose a noun phrase which typically has a certain relation with the verb, and can form a plausible script together with the verb. In this case the relation is locative, as generally “restaurant” indicates a place in relation to the verb “to eat.” In the thematic role violations (e.g., Fred eats a sandwich…, Fred eats in a restaurant…) we exchanged the noun phrases “sandwich” and “restaurant.” On the basis of the syntactic structure “sandwich” gets assigned a locative role, while “restaurant” is assigned a patient

**Table 1** Example of one quartet of the thematic role violations and their baseline in the original Dutch version and their English translation (depicted in italics).

| Condition                  | Structure | Sentence                                      |
|----------------------------|-----------|-----------------------------------------------|
| Baseline                   | NP–V–NP   | Fred eet een boterham…                        |
|                            | NP–V–PP   | (Fred eats a sandwich…)                       |
| Thematic role violation    | NP–V–NP   | Fred eet een restaurant…                     |
|                            | NP–V–PP   | (Fred eats a restaurant…)                    |
|                            |           | (Fred eats a sandwich…)                      |

In the thematic role violations the incrementally earlier syntactic constraints impose a certain thematic role onto the critical word, which conflicts with the plausible semantic-thematic role. The critical words are depicted in bold.
role. This conflicts with the semantic–thematic bias. Thus, while the content words (e.g., eat – sandwich – restaurant) are exactly the same for the baseline and thematic role violations and can form a plausible combination of words or a coherent script, the syntactic cues regarding the phrasal type (NP, PP) preceding the critical word differ. It is because of these, incrementally earlier, syntactic cues that the critical word stands in a certain relation with regard to the verb (e.g., in a sandwich), and this conflicts with the semantic–thematic bias.

A critical difference between our materials and those of for instance Kim and Osterhout (2005, The pizza had been delivering…) or Kuperberg et al. (2003, For every breakfast the eggs would eat…) lies thus in the relative strength of the syntactic constraints. Our conflicts between semantics and syntax concern not a local morpho-syntactic inflection, but it affects the head and thus the type of the phrase. Instead of merely changing the morpho-syntactic inflection, phrasal retyping – from NP to PP or from PP to NP – would be necessary to resolve the conflict at the semantic level. Therefore our materials contain more constraining syntactic cues as opposed to earlier experiments (Kuperberg et al., 2003; Kim and Osterhout, 2005).

Consequently the processing competition account predicts an N400 effect to the conflicts between semantics and syntax in our experiment (Father eats a restaurant… or Father eats in a sandwich…). As unambiguous syntactic cues regarding the phrasal type precede the critical word, the conflict between semantics and syntax lead to a higher processing load at the semantic level and thus elicit an N400 effect, despite the fact that the content words can be combined in a plausible way.

MATERIALS AND METHODS

PARTICIPANTS

Thirty-three native speakers of Dutch participated in the experiment, 24 of whom were included in the final analysis (12 males; age: 18–25 years, mean age 20.6). Participants were recruited from the Donders Institute subject pool. All had normal or corrected-to-normal vision and were right handed. None of the participants had any neurological or language impairment, nor had any of the participants participated in the pretests (see below). All participants signed informed consent and received reimbursement or course credits for participation. The study was approved by the local ethics committee. Nine of the participants were excluded from final analysis due to an excessive number of artifacts in the EEG signal.

STIMULUS MATERIAL

Control condition: semantic manipulation

We selected 80 Dutch sentence pairs containing a semantic violation and a correct control. These sentence pairs have already been used in other experiments and are known to elicit an N400 effect (Swaab et al., 1997; van den Brink et al., 2001; Hagoort et al., 2004). These sentences were included to serve as a basis for comparison of ERP effects to other experimental sentences within the same group of participants.

The experimental sentence pairs were identical with the exception of one word, which was the critical word for our analyses (printed in bold). Each pair consisted of a sentence that was completely semantically coherent (SC: Whipped cream tastes sweet and creamy) and a sentence that contained a semantic anomaly (SA: Whipped cream tastes anxious and creamy). The critical words were never in sentence-final position and were matched across conditions for word frequency (SC = 2.964, SA = 2.862), based on log lemma frequencies of the Dutch database CELEX (Baaijen et al., 1993), and length (SC = 5.69, SA = 5.73). The length of the sentences ranged from 5 to 19 words. The average length was 12.7 words (SD = 3.0).

Control condition: agreement manipulation

We also selected sentence pairs from Hagoort et al. (1993), where one sentence contained an agreement violation between the subject and the verb and the other served as a correct control. These agreement violations are known to elicit a standard P600 effect (Hagoort et al., 1993) and were also included to serve as a basis for comparison with ERP effects to other experimental sentences within the same group of participants. Again, the sentence pairs were equal with the exception of one word, which was the critical word for the ERP analysis. In half of the cases the critical word was the verb of the sentence (The spoiled child throws/throw the toys on the floor), in the other half the subject was the critical word (With an apple in the hand walk/walks the sisters to school; in Dutch the verb can appear in front of the subject). The length of the sentences ranged from 5 to 14 words. The average length was 10.7 words (SD = 1.69).

Thematic role manipulation

The experimental materials to investigate conflicts between thematic and grammatical roles with temporal precedence of grammatical cues contained 104 sets of four sentences in Dutch. An example is shown in Table 1. Each sentence started with an animate subject and a transitive verb for which the use of a direct object is optional. Two different sentence structures served as a baseline. In the first sentence (NP–V–NP; Fred eats a sandwich…) the verb takes a direct object. This argument consisted of a noun phrase which typically is a patient in relation to the verb. For instance, a “sandwich” is usually the patient of the verb “to eat.” In the second sentence the verb takes a prepositional object as argument (NP–V–PP; Fred eats in a restaurant…). Again we chose a noun phrase which typically has a particular relation with the verb. In this case this is a locative relation. In the thematic role violations we exchanged the noun phrases “sandwich” and “restaurant.” On the basis of the syntactic structure “sandwich” gets assigned a locative role, while “restaurant” is assigned a patient role. This conflicts with the semantic–thematic bias.

All critical words referred to inanimate entities and never occurred in sentence-final position. In addition, the baseline condition and thematic role violations contain exactly the same critical words. As a consequence the critical words of the thematic role violation are completely matched to the critical words in the baseline condition. The thematic roles of the direct objects in the baseline condition were patient (entity undergoing the effect of some action, often undergoing some change in state; Saeed, 1997), e.g., The chef cuts onions…, theme (entity which is moved by an action or whose location is described; Saeed, 1997), e.g., The field hockey player hits the ball…, or goal, e.g., The workmen build villas…. The thematic roles of the nouns in the prepositional phrases in the baseline condition could mostly be classified as
locative, e.g., John cooks in the **kitchen**…, instrumental, e.g., The maid sweeps with a **broom**…, or temporal, e.g., The witness spoke during the **hearing**… The word which followed the critical word was the same for every sentence within an item. Sentences varied in length from 5 to 14 words, with the average sentence length being 8.8 words (SD = 1.6).

A pretest with 63 participants was performed on these materials to assess whether the congruent (Fred eats a **sandwich**; Fred eats in a **restaurant**) and incongruent sentences (Fred eats a **sandwich**; Fred eats a **restaurant**) were rated accordingly. The sentences were presented up to and including the critical word. We omitted the rest of the sentences to obtain a judgment about the sentence at the place of the critical word without the influence of the lexical content of the rest of the clause. Because some of the sentences required a completion at this point of the sentence, we asked the participants to complete the sentences and to rate how difficult this was on a scale ranging from 1 “easy” to 5 “impossible.” On the basis of this pretest we selected 104 items for which the congruent sentences were easier to complete (mean = 1.44, SD = 0.44) than the incongruent sentences [mean = 3.98, SD = 0.62; \( t(414) = -49.15, p < 0.001 \)].

**Pretest contextual constraint semantic and thematic role manipulation**

We conducted an additional behavioral test to measure contextual constraint values of the correct controls for the semantic anomalies and thematic role violations. Sentence constraint was operationalized via cloze probability, i.e., the probability of production of the critical word of the correct control condition. Thirty-six participants, who had not participated in the previously mentioned pretest, were asked to complete the sentences. Sentences were presented up to the critical word. The average contextual constraint of the correct sentences of the semantic manipulation was higher (0.64, SD = 0.30) than that of the correct controls of the thematic manipulation [mean = 0.22, SD = 0.26; \( t(286) = 11.527, p < 0.001 \); see **Figure 4A**].

**Other materials**

We included 50 coherent items in the experiment as filler sentences. These sentences were selected from the Dutch CLEF corpus (van der Beek et al., 2001). Also, for purposes beyond the scope of this paper, a set of 120 syntactically ambiguous sentences was included. Additionally we included 20 practice-items, which were similar in nature to the experimental items.

**Experimental versions**

The materials were split into 12 different lists, with each list consisting of 454, pseudo randomly mixed sentences. This was done in such a way that each version of each item was distributed equally over these 12 lists, that all lists contained an equal number of items per condition, that no participant read the same sentence in more than one variant, and that each variant was read by an equal number of participants. Critical words were only used once in the critical position. If repetition of the critical word could not be avoided, use in the critical position preceded use in non-critical positions. The length of the sentences ranged from 5 to 19 words. The average length was 10.3 words (SD = 2.26).

**Procedure**

Participants were tested individually in a sound-attenuating booth. The booth was dimly lit (Fiber optic lights DMX 512 at 40%). Participants were seated in a comfortable chair and were told that the aim of the experiment was to investigate how people process sentences and that some of the sentences would be more difficult or stranger than other sentences. Participants were informed that they were going to see a written sentence that would be presented word-by-word in the middle of the computer screen. Participants were instructed to read the sentences carefully and to attempt to understand them as well as possible. They were asked to try not to move or blink during the presentation of the sentence. No other task demands were imposed.

After a short practice session, 434 trials were presented in five blocks of 15 min each, separated by rest periods of approximately 5 min each. Halfway through every block there was an additional 30 s break. Viewing distance was approximately 110 cm. The first word of the sentence started with a capital letter, the rest of the words were presented in white lowercase ARIAL (23-point font size) against a dark background in the center of a CFT 60 Hz computer screen. Each word was presented for 300 ms followed by a blank screen for 300 ms. The final word of the sentence ended with a period. After the final word an asterisk appeared for 2 s, indicating to the participants that they could blink and move their eyes. There was a 1.2-s blank interval between the asterisk and the start of the next trial. Sentences were presented using Presentation software (Neurobehavioral systems, www.neurobs.com).

**EEG RECORDING AND ANALYSIS**

The EEG was recorded from 28 cap-mounted Ag/AgCl electrodes (Easycap). Four electrodes were placed over the standard 10% system midline sites Fz, FCz, Cz, and Pz. Eleven pairs were located over the standard lateral sites FP1/FP2, F7/F8, F3/F4, FC5/FC6, FC1/FC2, T7/T8, C3/C4, CP5/CP6, CP1/CP2, P7/P8, and O1/O2. Two electrodes were placed at the outer left and right canthi to monitor horizontal eye movements. Vertical eye movements were monitored using FP1 and an electrode placed below the left eye. An additional electrode was placed on the right mastoid bone. During measurement, all electrodes were referenced to the left mastoid. Electrode impedances of the EEG- and EOG-electrodes were kept below 5 and 10 kΩ respectively. Signals were recorded with a BrainAmp DC amplifier (Brain Products, Germany), using a 125 Hz low-pass filter, a time constant of 10 s, and a 500-Hz sampling frequency. The software package Brain Vision Analyzer (Brain Products, Germany) was used to analyze the waveforms.

Offline, the EEG electrodes were rereferenced to the mean of the right and left mastoid and the EOG electrodes were converted into bipolar horizontal and vertical EOG signals. A 30-Hz, 12 dB low-pass, Hanning filter was applied. Subsequently, the critical words were segmented using a window which started 200 ms before and ended 1500 ms after the critical word. After baseline correcting to the 150 ms pre-critical-word interval, segments were semi-automatically screened for eye movements, electrode drifting, amplifier blocking and EOG artifacts using a 75 μV criterion. Segments containing such artifacts were rejected (13.9% overall, with no asymmetry across conditions). The remaining EEG segments were averaged per participant and per condition.
Nine participants were excluded from the analysis because more than 20% of the trials were rejected, leaving 24 participants for subsequent analysis.

A standard latency window of 300–500 ms after onset of the critical word was used to compute the mean amplitude of the N400 components, and a 500–900 ms latency window was applied for assessment of the P600 effect. The effects were evaluated in repeated measures analyses of variance (ANOVA). The effects were assessed using four quadrants, which consisted of the mean of five electrodes: left anterior (F3, F7, FC1, FC5, C3), right anterior (F4, F8, FC2, FC6, C4), left posterior (CP1, CP5, P3, P7, O1), and right posterior (CP2, CP6, P4, P8, O2). Interactions with the factor Quadrant were followed by single quadrant analyses. Separate ANOVAs were conducted for the midline sites (Fz, FCz, Cz, Pz). Greenhouse–Geisser correction was applied when evaluating effects with more than one degree of freedom in the numerator. In these cases the corrected $p$-values with the original degrees of freedom will be reported.

**RESULTS**

**CONTROL CONDITION: SEMANTIC MANIPULATION**

Figure 1A shows the grand-average waveforms of the semantic anomalies and the correct controls at Pz, as well as the topographical distribution of the effect of the semantic manipulation. These anomalies showed a clear N400 effect (for an overview of the statistical analyses see Table 2): the repeated measures ANOVA with the factors Semantic Fit and Quadrant in the 300–500 ms latency window revealed a main effect of Semantic Fit and a significant interaction between Semantic Fit and Quadrant. A main effect of Semantic Fit was found for every separate quadrant. The midline

![Figure 1A](image-url)

![Figure 1B](image-url)
analysis revealed a main effect of Semantic Fit as well. Even though the N400 effect is widespread, the topographical distribution shows that the effect is strongest over the posterior electrodes, which is common for N400 effects (Kutas and Van Petten, 1994).

CONTROL CONDITION: AGREEMENT MANIPULATION

Figure 1B shows the grand-average waveforms of the agreement violations and their correct controls at Pz, as well as the topographical distribution of the agreement effect. Table 3 shows an overview of the statistics. The repeated measures ANOVA with the factors Agreement and Quadrant within the 500–900 ms latency window revealed a marginal effect for Agreement, but a significant effect for the Agreement × Quadrant interaction. Subsequent tests in the anterior quadrants showed no main effects for Agreement. There were, however, main effects of Agreement in the posterior quadrants, revealing the posterior distribution of the P600 effect. The midline analysis revealed a marginal effect for Agreement.

The positivity of the waveforms of the agreement violations was long lasting. This had been confirmed by repeated measures ANOVAs in the 900–1200 ms time window. The repeated measures ANOVA with the factors Agreement and Quadrant showed a main effect for Agreement and an interaction effect between Agreement and Quadrant. Subsequent tests for the separate quadrants revealed the same pattern as in the earlier time window: the tests for the anterior quadrants showed no effects for Agreement, while the agreement violations were significantly more positive compared to the baseline condition in the posterior quadrants. Also the midline analysis revealed a significant main effect of Agreement.

THEMATIC ROLE MANIPULATION

The grand-average waveforms of the thematic role violations and the correct controls are presented in Figure 2A. After the N1–P2-complex the waveforms started to diverge around 300 ms after onset of the critical word. The waveforms of the thematic role violations are more negative than their correct controls. This negativity seems to be maximal at posterior electrodes and extends toward more frontal electrodes2.

The repeated measures ANOVA with the factors Thematic Fit and Quadrant within the 300–500 ms latency window after critical-word-onset revealed a main effect of Thematic Fit. There was no interaction between Thematic Fit and Quadrant. The midline analysis within the same time window also showed a main effect for Thematic Fit (see Table 4).

Figure 2B depicts the scalp distribution of the effect of Thematic Fit in the 300–500 ms time window. Even though the negative effect of the thematic manipulation is widespread over all electrodes, the effect is most pronounced at the posterior electrodes.

To verify whether the effect was comparable for the two sentence structures (NP–V–NP and NP–V–PP) we performed additional repeated measures ANOVAs. First with the factors Thematic Fit, Quadrant and Sentence Type within the 300–500 ms time window. There was no interaction between Thematic Fit and Sentence Type [F(1,19) = 2.465, MSE = 34.197, p = 0.133], nor was there an interaction between Thematic Fit, Sentence Type and Quadrant [F(3,57) < 1]. The second repeated measures ANOVA was performed on the midline electrodes with the factors Thematic Fit and Sentence Type. Here also the interaction between Thematic Fit and Sentence Type failed to reach significance [F(1, 20) = 2.503, MSE = 15.041, p = 0.129]. Hence, the two sentence types did not yield different effects.

COMPARISON THEMATIC ROLE AND SEMANTIC ANOMALIES

In Figure 3 the effects of the thematic role violations and the semantic anomalies are presented. With a mean effect size of −2.9 across all quadrants the semantic anomaly effect is approximately four

2Closer inspection of this frontal negativity seems to indicate that this results from items in the later part of the experiment.
The obvious difference in effect size still needs to be accounted for. A behavioral pretest (see Materials and Methods section) indicated that contextual constraint was higher for the correct controls of the semantic anomalies as compared to the correct controls of the thematic role violations. Subsequently, we computed the correlation between the contextual constraint score of an item and N400 effect size averaged over subjects for each electrode by means of computing the Pearson’s correlation coefficients. These correlation analyses showed that there were significant correlations between contextual constraint values of the correct-control-items and the ERP effects of the semantic and thematic role violations per electrode within the 300–500 ms time window [e.g., $r(Pz) = -0.42, p < 0.01$], indicating that the stronger a sentence frame biased in a certain direction, the larger the N400 effect size.

**Figure 4** displays the mean contextual constraint, the correlations between contextual constraint and the ERP effects at all the electrodes, as well as the scatter diagram and best-fitting regression line of electrode Pz across both sets of stimuli.

The topographical distribution and latency of the thematic role violation was comparable to the standard semantic manipulation (Whipped cream tastes sweat.../Whipped cream tastes times larger than the effect of the thematic role violations, which has a mean effect size of −0.8 across all four quadrants [Violation Type × Incongruency: $F(3, 69) = 26.865, MSE = 20.475, p < 0.001$].

For visual comparison of the topographical distributions of the two effects without interference of differences in effect size, we differentially scaled the topoplots. After this adjustment the semantic and thematic role effects reveal similar topographical distributions. This has been confirmed by an additional analysis of variance on scaled data within the 300 and 500 ms latency window. First, for every participant difference scores between the two conditions of the semantic and thematic role manipulations were computed for every electrode. Second, a scaling procedure was performed to avoid that amplitude differences between the semantic and thematic role effect would be incorrectly be interpreted as distribution effects.

In this procedure, the electrode-specific difference scores were $z$-transformed for each effect type separately (van den Brink et al., 2001). The $z$-transformed values were entered into an ANOVA that crossed the 2-level factor violation type with the 26-level factor electrode. This analysis revealed that the scalp distribution of the two effects did not differ significantly [Violation Type × Electrode: $F(5, 575) = 0.126, MSE = 4.899, p = 0.987$].

**Discussion**

We conducted an ERP study to investigate whether conflicts between semantics and syntax elicit N400 effects. Participants were presented with Dutch sentences containing such conflicts. In the baseline condition syntactic structure was congruent with the semantic–thematic bias, or the most plausible scenario (Fred eats a sandwich.../Fred eats in a restaurant...), while in the thematic role violations (Fred eats a restaurant.../in a sandwich...), the syntactic structure conflicted with the semantic–thematic bias. The thematic role violations elicited an N400 effect compared to the baseline condition.

The thematic role violation was comparable to the standard semantic manipulation (Whipped cream tastes sweat.../Whipped cream tastes
Kos et al. Conflicts between semantics and syntax

The finding of an N400 effect to sentences containing conflicts between semantics and syntax is in accordance with the hypothesis of the hypothesis of the processing competition account. The processing competition account assumes that the language processing system consists of highly interactive parallel processing streams (Jackendoff, 2002). Constraints of the independent systems are taken into consideration concurrently during listening or reading (Hagoort, 2003, 2005). Information from one level can influence the processing at another level. A conflict between semantic and syntactic constraints is signaled via the interface and the relative strength of the cues of the parallel processing streams determines the division of processing load (Hagoort et al., 2009). Early delivered cues take precedence over incrementally later delivered cues.

In the sentences “Fred eats a sandwich...” or “Fred eats in a restaurant...” the content words Fred – eat – sandwich or Fred – eat – restaurant can form a plausible scenario, just as in the sentences “For every breakfast the eggs would eat” (Kuperberg et al., 2003; Kim and Osterhout, 2005); or “the pizza has been delivering” (Kim

**FIGURE 3** Comparison of the effect of the thematic role and semantic manipulation: the grand-average waveforms at electrode Pz and the scalp topography with adjusted scales.

**Thematic role manipulation:**
Fred eats a **sandwich** / in a **restaurant**...
Fred eats a **restaurant** / in a **sandwich**...

**Semantic manipulation**
Whipped cream tastes **sweet**...
Whipped cream tastes **anxious**...

anxious...). However, N400 effect size of the thematic manipulation was smaller compared to the semantic manipulation. Two possible factors could explain the difference in effect size. The first is contextual constraint (Kutas et al., 2006), as a behavioral test revealed lower contextual constraint values for the materials of the thematic role manipulation compared to the materials of the semantic manipulation. Moreover, these contextual constraint values correlated with the N400 effects elicited by the thematic role and semantic manipulations.

In addition, the existing plausible lexical semantic association (Van Petten et al., 1997) between the critical word (sandwich or restaurant) and the verb (eat) may also have contributed to the smaller effect size of the N400 effect in the thematic role violations. It is known that lexical semantic relationships can influence the N400 amplitude independently from sentence-level semantic relationships (Van Petten et al., 1997). Thus, the amplitude difference within the N400 effect could in part be due to the presence of a lexical semantic association between “eat” and “sandwich” or “restaurant.” This lexical semantic association is lacking in the semantic anomalies.

**THEMATIC ROLE VIOLATIONS AND LANGUAGE PROCESSING MODELS**
The finding of an N400 effect to sentences containing conflicts between semantics and syntax is in accordance with the hypothesis of the processing competition account. The processing competition account assumes that the language processing system consists of highly interactive parallel processing streams (Jackendoff, 2002). Constraints of the independent systems are taken into consideration concurrently during listening or reading (Hagoort, 2003, 2005). Information from one level can influence the processing at another level. A conflict between semantic and syntactic constraints is signaled via the interface and the relative strength of the cues of the parallel processing streams determines the division of processing load (Hagoort et al., 2009). Early delivered cues take precedence over incrementally later delivered cues.

In the sentences “Fred eats a restaurant...” or “Fred eats in a sandwich...” the content words Fred – eat – sandwich or Fred – eat – restaurant can form a plausible scenario, just as in the sentences “For every breakfast the eggs would eat” (Kuperberg et al., 2003; Kim and Osterhout, 2005); or “the pizza has been delivering” (Kim
and Osterhout, 2005). In our sentences the moment at which the conflict takes effect was preceded by a clear syntactic cue regarding the phrasal type, overruling the plausible semantic relation between the critical word and the verb. This led to a higher processing load at the semantic level as reflected by the N400.

The processing competition theory can thus account for the pattern of findings for sentences in which plausible scenario's conflict with the syntactic structure. However, it seems less obvious how the processing competition account explains P600 effects to verbs in grammatically well-formed sentences in which there is no such plausible scenario: “For every breakfast the eggs would plant flowers in the garden...” (Kuperberg et al., 2007) or “The necklace is drinking the milk” (Nakano et al., 2010). In these sentences the verbs take an inanimate agent, while they require an animate agent. Animacy is a semantic cue. It could be that this semantic cue is immediately used in assigning thematic roles of the sentences. Experiments comparing inanimate with animate nouns at the beginning of the sentence show different ERP patterns to these words (Weckerly and Kutas, 1999; Nakano et al., 2010). It seems that subjects are sensitive to the fact that sentence initial inanimate nouns are not frequently used as agents (Bock, 1986). Consequently, participants might expect it to be a patient in a passive construction. As a result, the active voice of the verb is perceived as a syntactic error. It must, however be noted that not all animacy violations with respect to the verb lead to a P600 effect: In another study the verb in sentences such as “The dusty tabletops were devouring...” elicited an N400 effect. Individual variation (Nakano et al., 2010) or the presence of a more elaborate semantic context (Kuperberg, 2007; Kuperberg et al., 2007), which places more emphasis on semantics, might play a role here.

The results of the present study support the processing competition account, the question remains however, whether other models of language processing can explain our findings as well. The continued analysis account (Kuperberg, 2007) assumes that there are several independent, interactive processing streams: the semantic memory-based mechanism and combinatorial mechanisms: for instance semantic–thematic and syntactic streams. Several factors can trigger a conflict between these interactive streams: animacy violations, semantic associations between the arguments and the verb, an elaborate context, or presence of an acceptability judgment task. A conflict between these processing streams results in a prolonged combinatorial process, eliciting a P600 effect (Kuperberg, 2007).

The N400 finding for our conflicts (Fred eats in a sandwich...) fits with the continued analysis account. The close semantic relation between the content words is an insufficient trigger to elicit...
a prolonged combinatorial process. As the content words are the same across conditions, the N400 finding can be explained as a consequence of the interaction between semantic memory and the combinatorial mechanisms.

The Monitoring Theory and eADM encounter more difficulties explaining the N400 effect to our thematic role violations, due to the fact that in these models plausibility mechanisms operate independently from syntactic or role assignment operations. According to the Monitoring Theory, a syntactic algorithm and a semantic plausibility heuristic, which combines individual word meanings in the most plausible way, work in parallel and independently to get to the interpretation of the sentence (van Herten et al., 2006). In case of conflicting thematic output from the two processing streams, a monitoring process is triggered, eliciting a P600 effect. It is assumed that this monitoring process is general and not only triggered by conflicts between semantics and syntax, but also by conflicts between semantic and orthographical representations (Vissers et al., 2006) and by conflicts between semantic and conceptual representations (Vissers et al., 2006). Recently, it has been postulated that monitoring processes are only triggered by strong conflicts, as mild and strong conflicts elicited differential ERP responses. Mildly implausible sentences (the eye consisting of among other things a pupil, iris and eyebrow…) only elicited an N400 effect, whereas deeply implausible sentences (the eye consisting of among other things a pupil, iris and sticker…) elicited a biphasic N400–P600 pattern compared to plausible sentences (the eye consisting of among other things a pupil, iris and retina…). Within the Monitoring Theory it is assumed that in both conditions the implausibility leads to a semantic integration problem, which is reflected by the N400 effect. In the mildly implausible condition the integration difficulties are resolved, whereas in the deeply implausible condition integration fails and a monitoring process is elicited, as reflected by the presence of a P600 effect (van de Meerendonk et al., 2010).

The finding of an N400 effect for our conflicts (Fred eats in a sandwich…) challenges the assumption of the Monitoring Theory that the plausibility heuristic works independently from the syntactic algorithm. According to this assumption a conflict arises between the plausibility heuristic – sandwich is patient – and the syntactic algorithm (V–PP) in which no patient role is licensed. This conflict triggers a monitoring process and thus elicits a P600 effect. It might be argued that the conflict in our thematic role violations is not strong enough, and that the integration difficulties are, therefore, resolved at the semantic level, eliciting an N400 effect. However, the content words are the same across conditions and thus the integration using the plausibility heuristic should not be different.

According to the eADM (Bornkessel-Schlesewsky and Schlesewsky, 2008) the output of the assignment of generalized roles (actor, undergoer; compute prominence), the build-up of argument structure of the verb (compute linking) and the plausibility process is integrated within a generalized mapping step. In case of a conflict between these information sources, processing difficulties arise within the generalized mapping process, eliciting a P600 effect. Since the plausibility step cannot influence the prominence/linking operation, eADM predicts a P600 effect for our thematic role violations (e.g., Fred eats a restaurant…), as restaurant can get assigned the undergoer role without any processing difficulty and, therefore, no N400 effect should be expected. Also in the parallel plausibility step the absence of an N400 effect is predicted, since “eat” and “restaurant” forms a plausible combination of open class words. In the generalized mapping step the conflict between the role assignment and plausibility steps becomes apparent and, thus, a P600 effect would be predicted. In contrast to these predictions, an N400 effect was observed.

We must, however, emphasize that the eADM has no clear predictions regarding adjuncts (Fred eats in a sandwich…), as this model just focuses on arguments of the verb. It is only stated that these non-obligatory constituents are subject to different regularities compared to the processing of arguments (Bornkessel and Schlesewsky, 2006). The eADM needs to be elaborated to address these types of sentences properly.

In sum, our data can be explained by the fully interactive models (processing competition, semantic attraction, and continued analysis accounts), whereas the models which assume a relative independence of computing plausibility encounter problems accounting for our data.

QUANTIFICATION OF CUE STRENGTH, SUFFICIENCY OF TRIGGERS, AND CONFLICT STRENGTH

Three of the four language processing models specify conditions for how a conflict is processed, which are hard to objectively quantify. The processing competition account assumes that the relative strength of the syntactic and semantic cues determines which processing stream has to do the work to resolve the conflict. Plausibility ratings can give an indication for the strength of the semantic cues, assessing the strength of syntactic cues poses, however, a larger challenge. The continued analysis account has similar problems with the notion of sufficient amount of triggers, because it is hard to quantify exactly beforehand if sentences contain sufficient triggers or not. The Monitoring Theory also encounters the problem with the notion of conflict strength. Further research is needed to specify precisely which cases of opposing syntactic preferences and semantic–thematic biases result in N400 or P600 effects.

CONCLUSION

In this paper we investigated sentences in which grammatical constraints were in conflict with semantic–thematic biases. These conflicts with an incrementally earlier syntactic cue led to an N400 effect. These data are in accordance with the processing competition account, in which it is assumed that conflicts with relatively strong and incrementally earlier syntactic cues lead to a higher processing load at the semantic level.

ACKNOWLEDGMENTS

This research was supported by a SPINOZA prize from the Netherlands Organization for Scientific Research (NWO) to Peter Hagoort. We thank Tineke Snijders for comments on the manuscript.
