Examining the evidence for an independent semantic analyzer: An ERP study in Spanish

Clare Stroud\textsuperscript{1} & Colin Phillips\textsuperscript{1,2}

\textsuperscript{1}Department of Linguistics and \textsuperscript{2}Neuroscience and Cognitive Science Program

University of Maryland

Corresponding author:

Colin Phillips
Department of Linguistics
1401 Marie Mount Hall
University of Maryland
College Park, MD, 20742-7505

Telephone: (301) 405 3082
Fax: (301) 405 7104

Number of text pages: 50 (includes tables and figures)
Number of figures: 5
Number of tables: 6
ABSTRACT

Recent ERP findings challenge the widespread assumption that syntactic and semantic processes are tightly coupled. Syntactically well-formed sentences that are semantically anomalous due to thematic mismatches elicit a P600, the component standardly associated with syntactic anomaly. This ‘thematic P600’ effect has been attributed to detection of semantically plausible thematic relations that conflict with the surface syntactic structure of the sentence, implying a processing architecture with an independent semantic analyzer. A key finding is that the P600 is selectively sensitive to the presence of plausible verb-argument relations, and that otherwise an N400 is elicited (The hearty meal was devouring ... vs. The dusty tabletop was devouring ...: Kim & Osterhout, 2005). The current study investigates in Spanish whether the evidence for an independent semantic analyzer is better explained by a standard architecture that rapidly integrates multiple sources of lexical, syntactic, and semantic information. The study manipulated the presence of plausible thematic relations, and varied the choice of auxiliary between passive-biased fue and active-progressive biased estaba. Results show a late positivity that appeared as soon as comprehenders detected an improbable combination of subject animacy, auxiliary bias, or verb voice morphology. This effect appeared at the lexical verb in the fue conditions and at the auxiliary in the estaba conditions. The late positivity elicited by surface thematic anomalies was the same, regardless of the presence of a plausible non-surface interpretation, and no N400 effects were elicited. These findings do not implicate an independent semantic analyzer, and are compatible with standard language processing architectures.

Keywords: ERPs, syntax, semantics, thematic roles, Spanish, animacy, incremental parsing
1. Introduction

1.1. ERP Evidence for a Parallel Processing Architecture

It has been widely assumed in psycholinguistics and neurolinguistics that there is a tight coupling of semantic and syntactic processes (e.g., Ferreira & Clifton, 1986; Friederici, 2002; Hagoort, 2008; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994), and more specifically that semantic interpretations are built up by combining the meanings of words and phrases based upon the syntactic structure of a sentence. Furthermore, in linguistics it is a standard assumption that semantic composition of words and phrases is parasitic on the computation of syntactic structure (Heim & Kratzer, 1998; Pollard & Sag, 1994; Steedman, 2000). In terms of on-line comprehension, these assumptions imply that when an incoming word is encountered, syntactic information is used first to fit the word into the current structure, and then the semantic contribution of that word is added to the representation based on the syntactic role it has been assigned. This view has been challenged by recent electrophysiological findings that have been taken to show that semantic composition can proceed independently of syntactic structure, i.e., the language processor considers interpretations that are not licensed by the structure of the sentence (e.g., Hoeks, Stowe, & Doedens, 2004; Kim & Osterhout, 2005; Kolk, Chwilla, van Herten, & Oor, 2003; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). These findings have broad implications for the architecture of language processing, and thus merit close scrutiny.

Recent ERP studies in English (e.g., Kim & Osterhout, 2005; Kuperberg et al. 2003; Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007), Dutch (e.g., Kolk et al., 2003; van Herten, Chwilla, & Kolk, 2006; van Herten, Kolk, & Chwilla, 2005), and German (e.g., Friederici & Frisch, 2000) have elicited a P600 in response to semantic anomalies in simple and unambiguously grammatical sentences. This finding appears to directly contradict the traditional view of the N400 as an ERP component that is elicited by semantic anomalies (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Kutas &
Hillyard, 1980; Kutas & Hillyard, 1984) and the P600 as an ERP component that is classically elicited by syntactic anomalies (Friederici, Pfeifer & Hahne, 1993; Hagoort, Brown & Groothusen, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Holcomb, 1992). For example, relative to acceptable sentences such as (1a) and (1b), Kim and Osterhout (2005) recorded a P600 at the verb in sentences such as (1c). The critical word is italicized in all examples.

(1)    (a) The hearty meal was *devoured* by the kids.          (control)
        (b) The hungry boy was *devouring* the cookies.       (control)
        (c) The hearty meal was *devouring* the kids.        (P600)
        (d) The dusty tabletop was *devouring* the kids.     (N400)

In (1c), the NP *the hearty meal* is unambiguously the subject of an active sentence, and *hearty meals* do not devour things, therefore one might expect the processor to analyze this sentence as semantically anomalous. According to the traditional view of the N400 as a response associated with detection of a semantic anomaly, this anomaly should therefore elicit an N400. The fact that these sentences elicited a P600 instead of an N400 has been taken as evidence that the processor recognizes that *hearty meal* is a good theme argument but a poor agent argument for *devour*, despite the fact that the surface structure of the sentence does not license the theme interpretation. Kim and Osterhout suggested that this *semantic attraction* between the verb and the argument in (1c) was so strong that it led participants to analyze the sentence as syntactically ill-formed. In support of this interpretation, Kim and Osterhout showed that sentences in which the subject NP does not make a good theme for the verb, such as (1d), elicited an N400 rather than a P600. Since the surface forms of both (1c) and (1d) are semantically anomalous in the same fashion, the contrasting ERP responses to these sentences provides a key piece of evidence that the language processor is sensitive to surface-incompatible interpretations.
Although current accounts of this phenomenon reflect different underlying architectural approaches and disagree about the exact processes reflected by the P600 (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kim & Osterhout, 2005; Kuperberg, 2007; van Herten et al., 2005, 2006), they share the common view that the P600 is elicited when there is an inconsistency between the relations dictated by the structure of the sentence (hearty meal is in the subject position of an active sentence) and the most plausible thematic relations between the open class words (hearty meal is a good theme argument for devour). These accounts argue that the only way that the processor could be aware that this inconsistency exists is if it independently evaluates the plausibility of relations among neighboring open class words, while ignoring the structure of the sentence at least temporarily.

However, the elegant argument that can be constructed based on ERP responses to sentences like (1) deserves further attention. First, in an increasingly crowded set of findings on the ‘thematic P600’ there are few studies that present the clear contrast between semantically attractive and non-attractive verb-argument combinations seen in (1), and a number of findings that appear to be inconsistent with Kim and Osterhout’s generalization. Second, it is important to consider how well the findings fit with the widespread evidence in psycholinguistics for a language processor that rapidly integrates information from multiple sources (syntax, semantics, discourse, lexical probabilities) to continuously update its interpretation of an incoming sentence (e.g., Altmann & Steedman, 1988; Levy, 2008; MacDonald, Pearlmenter, & Seidenberg, 1994; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Accounts of the thematic P600 have paid relatively little attention to the role that sources of information other than the semantic relationship among open class words may play in guiding the interpretation pursued by the processor (but see Bornkessel-Schlesewsky & Schlesewsky, 2008). Many of the existing experimental results are compatible with the more standard architecture in psycholinguistics.

The current study uses evidence from Spanish to evaluate whether the thematic P600 effect motivates the addition of an independent, compositional semantic analyzer to the processing architecture, or whether this body of results is consistent with a single analyzer that
integrates multiple sources of information as soon as they become available. Specifically, the study manipulates several syntactic and semantic factors in the sentence to provide a better understanding of the conditions under which a P600 is elicited. The results suggest that the thematic P600 effect in sentences like (1) is not driven by consideration of an interpretation that is incompatible with the surface structure, but rather by features of the surface form of sentences.

1.2. Previous Findings on the ‘Thematic P600’

There are now many reports of P600s elicited by anomalies that do not appear to be syntactic in nature. Among these, studies that show a contrast between semantic anomalies that elicit a P600 and semantic anomalies that elicit an N400 are particularly relevant for motivating the addition of an independent, compositional semantic analyzer to the processing model, because they potentially shed light on the circumstances under which the different processing streams arrive at inconsistent analyses.

Four main accounts have been proposed for the thematic P600 (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kim & Osterhout, 2005; Kuperberg, 2007; van Herten et al., 2005, 2006). These accounts reflect a range of fundamental assumptions about the nature of the language processor and they focus on accounting for somewhat different sets of results. Nevertheless, all accounts share a common architectural assumption of a processing architecture in which multiple analyzers operate in parallel and partially independent of one another; one analyzer operates primarily over structural information, and another analyzer operates primarily over semantic information. The semantic analyzer operates over open class words, and uses them to construct interpretations with little regard to the structural relations among them. Meanwhile, the syntactic analyzer provides an analysis of the sentence that is compatible with the morphosyntactic features of the sentence. Kuperberg (2007) allows for the additional possibility that a separate analyzer calculates likely thematic relations based on a limited set of features such as animacy. According to these accounts, a P600 is elicited when there is inconsistency between the analyses pursued by the different analyzers, or when one analyzer successfully reaches an analysis but
another does not. The four accounts differ in their precise definition of the different analyzers, and in their view of the specific process(es) that elicit the P600. A summary of these accounts can be seen in Table 1.

**Table 1.** Summary of current accounts of the thematic P600.

<table>
<thead>
<tr>
<th>Processing Analyzers</th>
<th>The P600 reflects…</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kim &amp; Osterhout (2005)</td>
<td>Syntactic &amp; semantic</td>
</tr>
<tr>
<td>van Herten et al. (2005, 2006)</td>
<td>Algorithmic &amp; heuristic</td>
</tr>
<tr>
<td>Kuperberg (2007)</td>
<td>Combinatorial (thematic and morphosyntactic) &amp; semantic-association</td>
</tr>
<tr>
<td>Bornkessel-Schlesewsky &amp; Schlesewsky (2008)</td>
<td>Linking &amp; plausibility</td>
</tr>
</tbody>
</table>

A cornerstone in the argument for an independent, compositional semantic analyzer is Kim and Osterhout’s (2005) evidence that the P600 is selectively elicited in the presence of plausible thematic relations among open class words. Kim and Osterhout explicitly manipulated the plausibility of the semantic fit between the subject NP and the verb, and found that a P600 was elicited in the condition in which there was a plausible relationship between the subject and the verb (semantic attraction: 1c), whereas an N400 was elicited in the condition in which there was no plausible relationship between the noun and the verb (no attraction: 1d). (Figure 3 of Kim and Osterhout’s study suggests that there was a late positivity in the no-attraction condition, of similar amplitude to the N400, but this difference was not found to be reliable in the statistical analyses.) In the absence of evidence of this kind that shows that the thematic P600 effect is modulated by the plausibility of relations among open class words, findings about the thematic P600 do not motivate an independent, compositional semantic analyzer. They may instead be understood as showing that the P600 is sensitive to specific features of surface syntactic and semantic analyses, such as lexical association or selectional restrictions involving features such as animacy.

However, findings from other studies present a more nuanced picture, and a number of studies have not found the presence of a plausible thematic relationship among the open class
words to be the key factor in eliciting the P600. Van Herten and colleagues (2006, Experiment 2) manipulated the strength of semantic association between the verb and the noun in object position of Dutch sentences, while holding constant the plausibility of the relations among the open class words in the anomalous sentences. Relative to acceptable sentences such as (2a), anomalous sentences such as (2b) in which the object and the verb were highly associated elicited a robust P600. In contrast, anomalous sentences such as (2c) in which the object and the verb were not highly associated elicited a robust N400 and a greatly attenuated P600 that was small at left hemisphere electrodes and absent at right hemisphere sites.

(2)  

(a)  Jan zag dat de olifanten de bomen omduwden...  
     (control)  
     ‘John saw that the elephants pushed-over the trees…’

(b)  Jan zag dat de olifanten de bomen snoeiden...  
     (P600)  
     ‘John saw that the elephants pruned the trees…’

(c)  Jan zag dat de olifanten de bomen verwenden...  
     (N400)  
     ‘John saw that the elephants caressed the trees…’

Likewise, Kuperberg and colleagues elicited a P600 in sentences with and without a plausible thematic relationship between the inanimate subject noun and the verb (3c-d) (Kuperberg et al., 2007).

(3)  

(a)  Every morning at breakfast the boys would eat…  
     (control)

(b)  Every morning at breakfast the boys would plant…  
     (N400)

(c)  Every morning at breakfast the eggs would eat…  
     (P600)

(d)  Every morning at breakfast the eggs would plant…  
     (P600)
Van Herten and colleagues point out that although the focus of Kim and Osterhout’s study was to manipulate the plausibility of the subject NP as a theme of the verb, their materials also varied in terms of the degree of semantic association between the words in the different conditions, i.e., *meal* is more strongly associated with *devour* than is *tabletop*. Van Herten and colleagues explicitly manipulated both of those factors, and concluded that the critical factor for eliciting the P600 was the presence of semantic association between the open class words (van Herten et al., 2006). Similarly, Kuperberg and colleagues elicited a P600 in sentences in which there was semantic association but no plausible relationship among open class words (Kuperberg et al., 2006, 2007). However, because the items in these three studies differed structurally, it remains unclear whether the critical factor for eliciting the P600 is semantic association in all cases or whether the existence of a plausible thematic relationship is relevant in certain structural configurations but not in others.

The picture is further complicated by two other sets of findings. First, there are a number of recent reports of P600s that are not straightforwardly described in terms of thematic relations. For example, a P600 has also been elicited by orthographic errors in high-cloze situations but not in low-cloze situations (Vissers, Kolk & Chwilla, 2006), and when the meaning of the sentence fails to match an accompanying picture (Vissers, Kolk, van de Meerendonk, & Chwilla, 2008). Second, other studies have elicited both an N400 and a P600 to words that are highly incongruous with the context, relative to words that are only slightly incongruous, which elicited only an N400 (Geyer, Holcomb, Kuperberg, & Pearlmutter, 2006; van de Meerendonk, Kolk, Vissers, & Chwilla, 2008). Relatedly, it is important to note that many conventional ERP studies of semantic anomaly have elicited an N400 followed by a late positivity, although the later component has received much less attention than the N400 (e.g., Federmeier, Wlotko, de Ochoa-Dewald, & Kutas, 2007; Matsumoto, Iidaka, Haneda, Okada, & Sadato, 2005; Van Petten & Luka, 2006).
In addition to the questions raised by varying empirical findings about the thematic P600, two additional issues complicate the interpretation of evidence for an independent compositional semantic analyzer. The first concerns the way in which different sources of information are used to incrementally interpret sentences that elicit the thematic P600. The second concerns the functional significance of N400 effects and their contribution to the evidence for independent semantic interpretation.

Many discussions of the thematic P600 have focused on the syntactic and semantic relations that obtain at the point when all of the critical open-class words have been processed (e.g., hearty meal … devour), but have paid less attention to the sequence of interpretive steps that leads up to that point (for related observations see Bornkessel-Schlesewsky & Schlesewsky, 2008). As the sentence unfolds the processor may rapidly draw on multiple sources of information to build expectations about the interpretation of the sentence. In so doing, the processor may construct plausible interpretive hypotheses that turn out to conflict with the ultimate form of the sentence, and such conflicts could arise within a single analyzer that rapidly integrates multiple sources of information, rather than processing different information sources in parallel. For example, in the case of sentences like The hearty meal was devouring … the inanimate subject, the auxiliary, and the verb stem devour- are all most consistent with a passive construction. If comprehenders integrate the verb stem into their ongoing interpretation before they recognize the import of the progressive –ing morpheme, then a conflict may arise due to a mismatch between an initial passive interpretation that must then be retracted. This conflict could account for the appearance of a P600, without appeal to an independent semantic analyzer. The same conflict would not arise in a sentence like The dusty tabletop was devouring … because the combination of the subject noun and the verb stem fails to support a passive analysis, even before the progressive verb morphology is processed. Consequently, the evidence from these sentences for independent semantic interpretation depends on the relative timing with which the processor uses information provided by the verb stem and the verb suffix.
A second interpretive complication arises because of ongoing debate about the interpretation of the N400, which makes it less clear whether the presence or absence of an N400 is informative in the question of an independent semantic analyzer. Under the integration view the amplitude of the N400 reflects the ease or difficulty of the semantic integration of an incoming word into the previous context (Baggio, Choma, van Lambalgen, & Hagoort, 2009; Brown & Hagoort, 1993; Hagoort, 2008; Kutas & Hillyard, 1980; Osterhout & Holcomb, 1992; van Berkum, Hagoort, & Brown, 1999). Under this view, the lack of an N400 in response to *the hearty meal was devouring*… supports Kim and Osterhout’s contention that semantic composition proceeds independently of structural information, because it suggests that the incoming verb root *devour-* is smoothly integrated into the preceding semantic context. In contrast, under the lexical view the amplitude of the N400 reflects the ease or difficulty with which the incoming word can be accessed from memory (Deacon, Hewitt, Yang, & Nagata, 2000; Federmeier, 2007; Fischler, Bloom, Childers, Roucos, & Perry, 1983; Kutas & Federmeier, 2000; Lau et al., 2008; van Berkum, in press). According to this view the N400 does not directly reflect combinatory processes, although factors such as semantic association and predictability based on the context influence the ease with which the incoming word is accessed. Thus, the lack of an N400 in the *hearty meal was devouring* … does not indicate that semantic integration has proceeded smoothly, but rather reflects strong lexical association between the noun and the verb or the predictability of the semantic features of the verb when it follows the subject noun.

In sum, there are various possible accounts of the thematic P600 that do not entail an independent semantic analyzer. In particular, two alternative interpretations of previous findings could weaken the evidence for an independent semantic analyzer. First, the thematic P600 in sentences like *The hearty meal was devouring* … may reflect consideration of a passive analysis that was fully compatible with the surface structure of the sentence until the point when the processor recognized the import of the progressive –*ing* suffix. Second, the contrasting ERPs elicited by semantically attractive and non-attractive verb-argument combinations may reflect
differences in lexical association rather than differences in the plausibility of the relationship between the noun and the verb. The current study was designed to address both of these interpretive concerns.

The current experiment

The aim of the current study was to investigate in greater detail the circumstances in which the P600 component is elicited by syntactically well-formed sentences with thematic anomalies. The ultimate aim of this research is to clarify whether findings of thematic P600 effects motivate the existence of an independent, compositional semantic analyzer that ignores syntactic cues, or whether these findings are also compatible with a single analyzer that rapidly integrates diverse sources of information (syntactic, semantic, contextual) as soon as they become available. The strongest evidence for an independent, compositional semantic analyzer would consist of experimental results showing that (a) the language processor identifies plausible thematic relations despite contradictory syntactic information that is available prior to the semantic composition process, and that (b) the thematic P600 effect is specifically sensitive to the plausibility of thematic relations rather than to lower-level lexical-semantic association effects. In this study we present thematic anomalies in environments that provide both of these circumstances, and thus test the extent to which the thematic P600 reflects syntax-independent semantic attraction between the words of a sentence.

We created these two circumstances by using an experimental design in Spanish that added two key features to the manipulation of thematic fit between a subject noun and a verb used in previous studies (e.g., Kim & Osterhout, 2005, Experiment 2). First, we varied the choice of auxiliary verb in order to manipulate the syntactic bias of the sentence towards an active or passive continuation of the sentence. This allowed us to establish stronger syntactic biases before readers encountered potentially contradictory thematic information. Second, we introduced a context sentence prior to each target sentence, which established all critical nouns as discourse referents in advance of the target sentence, thereby reducing the possible confound between
lexical predictability and thematic plausibility found in previous studies and allowing us to more directly assess the contribution of thematic plausibility to the P600.

In English, the auxiliary *be* is used in both passive and active progressive constructions. As already discussed, this leads to uncertainty in the interpretation of P600 responses elicited by semantically anomalous progressive constructions such as *The hearty meal was devouring*.... Thus, if comprehenders show sensitivity to the possibility that the subject NP is a good theme for the verb *devour* this could reflect either an independent semantic analysis that ignores the surface syntax, or a strongly incremental and non-independent analysis that pursues the possibility of a passive parse until it recognizes that the progressive –*ing* suffix blocks this analysis. In Spanish, different auxiliaries typically precede verbs in passive and active progressive constructions, and hence the choice of auxiliary can be used to manipulate syntactic expectations in advance of the thematic information conveyed by the main verb. The auxiliary *ser* provides evidence that an active progressive verb form is unlikely to follow, and the auxiliary *estar* provides evidence that a passive verb form is unlikely to follow.

In a design similar to Kim & Osterhout’s (2005) Experiment 2, felicitous control sentences were compared with semantically anomalous sentences with semantically attractive and non-attractive thematic relations. In the semantic-attraction condition the subject NP was a plausible theme for the verb, whereas in the no-attraction condition the subject NP was not a plausible theme for the verb. The three conditions (control, semantic-attraction and no-attraction) were tested with each of the two auxiliaries. The auxiliaries *ser* and *estar* were used in 3rd person singular past forms, *fue* and *estaba* respectively. *Fue* is typically used in passive constructions, and *estaba* typically is used in active progressive constructions. It is important for the current design that these are probabilistic biases: *fue* may be followed by a verb in progressive form, and *estaba* may appear in a passive construction. In addition, both auxiliaries can also be followed by an adjective. In the terms of Carlson (1977) the auxiliary *ser* is used with individual-level predicates that denote more permanent properties, such as *tall* or *intelligent*, and the auxiliary
**est**ar is used with stage-level predicates that denote more transient properties, such as *sick* or *hungry*. An example set of materials is shown in Table 2.

**Table 2.** Sample materials set.

<table>
<thead>
<tr>
<th>Primary Conditions</th>
<th></th>
</tr>
</thead>
</table>
| Context sentence   | Después del fuego, el bombero formuló un aviso para prohibir la entrada al piso quemado.  
*After the fire, the fireman prepared a warning to forbid entrance into the burnt apartment.* |
| **Fue** - passive prediction |  |
| Grammatical control | El aviso fue declarado por el bombero ante todos los vecinos.  
*The warning was declared by the fireman in front of all the neighbors.* |
| Semantic attraction | El aviso fue declarando con un megáfono para que todo el mundo lo oyera.  
*The warning was declaring with a megaphone so everyone would hear it.* |
| No attraction | El piso fue declarando con un megáfono para que todo el mundo lo oyera.  
*The apartment was declaring with a megaphone so everyone would hear it.* |
| **Estaba** - progressive prediction |  |
| Grammatical control | El bombero estaba declarando el aviso cuando el fuego se formó otra vez.  
*The fireman was declaring the warning when the fire rose up again.* |
| Semantic attraction | El aviso estaba declarando con un megáfono para que todo el mundo lo oyera.  
*The warning was declaring with a megaphone so everyone would hear it.* |
| No attraction | El piso estaba declarando con un megáfono para que todo el mundo lo oyera.  
*The apartment was declaring with a megaphone so everyone would hear it.* |

**Semantic Control Sub-Experiment**

| Context sentence | Cuando Juan se hizo millonario, contrató a un chófer y compró un carro costoso.  
*When Juan became a millionaire, he hired a chauffeur and bought an expensive car.* |
| Semantic Felicity |  |
| +Felicitous | El chófer estaba limpiando el carro *lujoso*, cuando Juan chocó con una moto contra la puerta del garaje.  
*The chauffeur was cleaning the luxurious car, when Juan crashed his motorcycle into the garage door.* |
| -Felicitous | El chófer estaba limpiando el carro *enfadado*, cuando Juan chocó con una moto contra la puerta del garaje.  
*The chauffeur was cleaning the angry car, when Juan crashed his motorcycle into the garage door.* |

**Syntactic Control Sub-Experiment**

| Context sentence | Cuando el ladrón entró por la ventana de la casa, la alarma antirrobo empezó a sonar muy alto.  
*When the thief entered the house through the window, the burglar alarm began ringing loudly.* |
| Syntactic agreement |  |
| +Agree | El ladrón estaba desconectando los *hilos* de la alarma antirrobo cuando llegó la policía.  
*The thief was disconnecting the(pl.) wires of the burglar alarm when the police arrived.* |
| -Agree | El ladrón estaba desconectando el *hilo* de la alarma antirrobo cuando llegó la policía.  
*The thief was disconnecting the(sg.) wires of the burglar alarm when the police arrived.* |

A context sentence was presented before each target sentence, in order to increase the naturalness of the passive construction and to reduce the contribution of lexical accessibility and lexical expectations to observed ERP effects. Passive constructions in Spanish have a more limited distribution than their English counterparts, and in particular they are more natural when
the passive subject NP is an existing discourse referent. Since three different subject NPs were used across the 6 target conditions in each item set, the target sentences within each item set were preceded by an identical context sentence that introduced all three potential subject NPs. In order to construct a natural context sentence, all three NPs were generally chosen from the same semantic field. Following the context sentence, it was still the case that the thematic fit between the verb and the subject NP was stronger in the control and semantic-attraction conditions than in the no-attraction conditions, but the global semantic association between the verb and the three subject NPs was similar across conditions, because the subject NPs were drawn from the same semantic field and all corresponded to existing discourse referents. This contrasts with the materials in Kim and Osterhout’s study, in which the no-attraction conditions were formed by combining the subject NP from the semantic attraction condition in one item set with a verb and sentence ending from another item set, resulting in a very low level of semantic association between the subject NP and the verb in the non-attractive condition.

Predictions

We can articulate a number of different predictions for the ERP results, based on different possible processing architectures. A central question is whether the semantically anomalous sentences in our study elicit a thematic P600 effect, and whether this effect specifically requires semantic attraction between the subject NP and the verb. If a P600 is elicited by a broader range of noun-verb combinations, or only when prior context favors a passive analysis, this undermines a key piece of evidence for a semantic analyzer that generates interpretations independent of syntactic cues.

If the processing architecture includes an independent, compositional semantic analyzer, then the processor should recognize plausible thematic relationships independent of syntactic information that is incompatible with the particular thematic interpretation. Further, if the P600 reflects detection of a conflict between a plausible thematic relationship and the surface syntax of
a sentence, then a P600 should be elicited in the semantic-attraction conditions and an N400 should be elicited in the no-attraction conditions, independent of the choice of auxiliary.

In contrast, in a processing architecture where a single analyzer rapidly integrates multiple sources of information as soon as they are available, then the auxiliary manipulation in Spanish should bias semantic interpretations to the thematic relations that are compatible with the biases of the auxiliary that precedes the main verb. In the conditions with *fue* the processor should be biased towards interpretations that are consistent with a passive structure. In contrast, in the *estaba* conditions the processor should be biased towards non-passive interpretations. The specific predictions from this architecture depend further on whether the P600 reflects a conflict between an initially plausible thematic interpretation and the surface syntax of the sentence, or broader properties of the surface form of the sentence, such as lexical semantic association or animacy violations.

First, if the P600 reflects a conflict between an initially plausible thematic interpretation and the surface syntax, then our manipulations should elicit a P600 only in the condition where the auxiliary *fue* suggests a passive construction and semantic attraction between the subject NP and the verb supports a passive interpretation. The three other anomalous conditions should elicit a P600. Alternatively, if the P600 is not sensitive to plausible thematic relations, but instead reflects broader properties such as lexical semantic association or animacy violations, then we should expect to observe a P600 in all four anomalous conditions. This is because the current study drew the subject NPs in the semantic-attraction and no-attraction conditions from the same semantic field and introduced all potential subject NPs as discourse referents in the context sentence.

*Control sub-experiments*

Because relatively few ERP studies have been conducted in Spanish, and particularly in the Latin American Spanish used in our study, the experiment also included two control sub-experiments that were used to confirm that syntactic and semantic anomalies elicit canonical
N400 and P600 effects in our participants. These sub-experiments are potentially valuable in the interpretation of novel or unexpected ERP effects observed in the primary conditions. The semantic sub-experiment manipulated the felicity of the semantic relation between a noun and a following adjective, and the syntactic sub-experiment manipulated the correctness of number agreement between a determiner and a noun. In order to control for possible effects of number morphology (Wagers, Lau, & Phillips, 2009) half of the syntactically anomalous sentences involved a singular determiner/plural noun mismatch and the other half involved a plural determiner/singular noun mismatch. Likewise, half of the corresponding grammatically correct versions contained a plural determiner and nouns, and the other half contained a singular determiner and noun. The items of the control sub-experiments were intermixed with items from the primary study. All items in the control sub-experiments were preceded by a context sentence, so that they would be indistinguishable from the other sentences in the study. A sample set of items for the control sub-experiments is shown in Table 2.

The effect of semantic felicity between a noun and its adjective has been previously investigated in several Spanish ERP studies. Martín-Loeches and colleagues compared sentences containing felicitous noun-adjective combinations such as *el sentimiento profundo* (‘the profound sentiment’) with infelicitous noun-adjective combinations such as *el sentimiento peludo* (‘the furry sentiment’) and observed a clear N400 (and also, unexpectedly, a small P600) (Martín-Loeches, Nigbur, Casado, Hohlfeld, & Sommer, 2006). In a series of auditorily and visually presented mini-stories Wicha and colleagues compared Spanish counterparts of *Little Red Riding Hood was carrying food to her grandmother in a very pretty {basket/crown}*{\textsuperscript{,}} which also elicited an N400 at the infelicitous word (Wicha, 2003; Wicha, Moreno, & Kutas, 2003). These experiments indicate that semantic anomalies elicit an N400 in Spanish in a similar fashion to other languages.

The syntactic control sub-experiment tested the effects of number (dis)agreement between an article and a noun. Previous studies of grammatical agreement in Spanish have yielded comparable results to similar manipulations in English. Previous studies that manipulated
either number or gender agreement between nouns and following adjectives have observed a P600 response and in most instances also a LAN (Barber & Carreiras, 2005; Martín-Loeches et al., 2006; Wicha, Moreno, & Kutas, 2004). In studies using a slightly different paradigm Barber and Carreiras (2003; 2005 experiment 1) presented word pairs (article plus noun or noun plus adjective) that disagreed in number, and elicited an N400. This finding is surprising in comparison to previous studies of agreement mismatches in English and Spanish, but plausibly reflects differences between the processing of isolated word pairs and words in sentence contexts.

2. Methods

2.1 Participants

There were 31 participants in the ERP study. Data from 1 participant were excluded due to technical problems; 5 participants were excluded due to low rates of accuracy in the behavioral task (< 75%), and 1 participant was excluded due to high levels of artifacts in the EEG recordings. All 24 remaining participants (14 female; mean age 27; range 18-41 years) were healthy, native speakers of Latin-American Spanish (from Perú (5), Argentina (4), Chile (4), Colombia (4), Puerto Rico (2), Guatemala (2), Venezuela (1), Ecuador (1), Costa Rica (1)). All had been in the United States for three years or less (mean: 1.5 years; range: 1 month to 3 years) except for one participant who had been in the country for five years but spoke Spanish almost exclusively in everyday life. All were right-handed and had normal or corrected-to-normal vision. All participants gave informed consent and were paid $15/hour for their participation, which lasted approximately 4 hours, including set-up time.

2.2 Materials

The primary materials consisted of sets of six sentences organized in a 2 × 3 factorial design, illustrated in Table 2. There were two levels of the auxiliary factor (fue and estaba) and three levels of the relatedness factor (control, semantic-attraction and no-attraction). The
The semantic control sub-experiment consisted of pairs of sentences containing felicitous and infelicitous noun-adjective combinations and the syntactic control sub-experiment consisted of pairs of sentences containing grammatical and ungrammatical article-noun agreement combinations. Examples of sentences from the semantic and syntactic control sub-experiments are shown in Table 2.

The target sentence in all six primary conditions began with a subject NP (determiner + noun), followed by an auxiliary (either *fue* or *estaba*), followed by a verb with either passive or progressive verbal morphology. Three conditions used the auxiliary *fue*, which is primarily used in passive constructions, and three conditions used the auxiliary *estaba*, which is primarily used in progressive constructions in Spanish. In the grammatical control condition with the auxiliary *fue* the sentence-initial NP was followed by a verb with passive morphology. In the grammatical control condition with the auxiliary *estaba* the sentence-initial NP was followed by a verb with progressive morphology. The anomalous semantic-attraction and no-attraction conditions contained verbs with progressive morphology, and differed in the goodness-of-fit of the inanimate subject NP as a theme for the verb.

In all conditions the target sentence was preceded by a context sentence, which was identical for all 6 conditions in each item set, and mentioned all three possible subject NPs of the target sentence. The three subject NPs of the target sentences were distributed as follows: an inanimate NP that was a good theme for the verb was used in the grammatical control condition with *fue* and in the semantic-attraction condition with both *fue* and *estaba*; an inanimate NP that was a poor theme for the verb was used in the no-attraction conditions with both *fue* and *estaba*; and an animate NP was used in the grammatical control condition with *estaba*. Although 5 of the 6 target sentences in each item set turned out to be active progressive constructions, the context sentences were designed to be compatible with both passive and active progressive continuations.

An important feature of the experimental design is that the two auxiliary verbs are biased to either active or passive voice constructions, but both also allow the other voice. *Ser* (and its
past tense form *fue*) is biased toward the passive voice, but can also be used in a progressive active construction (e.g., *fue buscando* which translates roughly as ‘went around looking for’). Likewise, *estar* (and its past tense form *estaba*) are biased toward the progressive active construction, but can also be used in a passive construction (e.g., *estaba cerrado por* ‘was in the state of being closed by’). If these continuations were impossible rather than improbable, then the anomalous sentences would contain straightforward syntactic violations, which would undermine the investigation of the factors that modulate the thematic P600.

In order to increase naturalness, the grammatical control conditions with *fue* and *estaba* used the voice that is most commonly associated with each auxiliary, with the result that the two grammatical control conditions differed in voice. The grammatical control condition with *fue* was a passive construction, the Spanish counterpart of sentences like *the hearty meal was devoured* …, and the grammatical control condition with *estaba* was an active progressive construction, the Spanish counterpart of *the hungry boy was devouring* …. These different constructions were chosen for the control conditions in order to ensure that each auxiliary was used in its preferred construction. Previous research suggests that using the two grammatical control conditions was appropriate. Kim & Osterhout (2005, Experiment 1) compared the anomalous *the hearty meal was devouring* to both an active control (*the hungry boy was devouring…*) and a passive control (*the hearty meal was devoured…*), and found that the choice of control did not impact the results.

For each set of six sentences, a verb and three subject NPs were chosen. The verb was chosen to be felicitous in both passive and progressive constructions, as judged by native Spanish speakers who assisted in the development of the materials. The verbs chosen for this experiment tended to impose less rigid selectional restrictions than those used in many previous studies of the thematic P600 (e.g., Kim & Osterhout, 2005; Kolk et al. 2003; Kuperberg et al. 2003, 2006, 2007; van Herten et al., 2005, 2006), due to the fact that few concrete verbs are felicitous in the passive voice in Spanish. The most felicitous verbs are typically those from business or newspaper language, such as *investigar* ‘investigate’, *fírm* ‘sign’, *entregar*
‘deliver’. A different verb was used in each item set, with the exception of 12 verbs that were repeated once each in order to increase the felicity of certain items. Three NPs were chosen as possible subjects for each verb: an animate noun that was a good agent for the verb, an inanimate noun that was a good theme for the verb, and another inanimate noun that was a poor theme for the verb. In order to ensure that the initial determiner in the target sentence did not provide a clue to which of the three previously-mentioned nouns was the subject of the target sentence, all three nouns in an item set shared the same grammatical gender. In order to guard against possible confounds in the ERP responses due to lexical differences based on factors such as word length or frequency almost all of the inanimate nouns (96%) were used in one item set as the semantically attractive noun and in another item set as the non-attractive noun. This was not the case for 4% of the item sets, in which the inanimate nouns were replaced in order to increase the felicity of the target sentences.

In order to guard against the possibility that participants might consistently assign the same interpretation to the violations, the material that followed the critical verb in the anomalous conditions of each item set was varied. This was important, because one account of previous findings about the thematic P600 claims that a P600 provides evidence that the processor entertained a passive interpretation to the sentence, which led to the perception that the progressive form was a syntactic error, and further claims that an N400 provides evidence that the processor considered an active analysis, which led to the perception of a semantic error, since the inanimate subject NPs was a poor agent for the verb (Kim & Osterhout, 2005). The material following the verb in the anomalous conditions might change the way that participants perceived the violation, and if this perception was reinforced across items, it could potentially change the ERP response that was evoked by the target verb. For example, if participants always saw a by-phrase after the anomalous inanimate NP + progressive form, it might lead them to favor a passive interpretation for the sentence as a whole, and therefore perceive the violation as a syntactic error, due to verbal morphology that mismatches the preferred interpretation. Therefore, the completions in the anomalous conditions were distributed equally among three
options: a by-phrase (e.g., ‘by the fireman’), an inanimate direct object NP (e.g., ‘the apartment’), or a preposition or other adverb (e.g., ‘with a megaphone’ or ‘many times’). In most item sets the word that followed the critical verb was short (i.e., prepositions and definite articles), with some slightly longer adverbs.

Materials for the two control sub-experiments began in a very similar form to the target experimental items. A context sentence was followed by a target sentence that began with a sequence of a determiner, a noun, an auxiliary, a verb, and then the post-verbal continuation. These items were designed to make it more difficult for participants to anticipate specific features of the target conditions. Across the entire set of experimental materials the target sentences contained equal numbers of the following beginnings: inanimate NP + fue, inanimate NP + estaba, animate NP + fue and animate NP + estaba.

180 sets of 6 items for the primary study were distributed across six presentation lists in a Latin Square design such that each list contained 30 items per condition. In addition, 60 pairs of items in the syntactic control sub-experiment were distributed across two presentation lists, and 60 pairs of items in the semantic control sub-experiment were distributed across two further presentation lists. The lists for the two control sub-experiments were then crossed such that there were four presentation lists for the control sub-experiments. The six lists of target items were crossed with the four lists of control sub-experiment items and 60 additional filler items to create a total of 24 lists with 360 items each. Each list was seen by one of the 24 participants included in the analysis. The filler items were similar to the experimental items in maintaining a two-sentence format, and were all grammatically correct but used a variety of different syntactic structures. Thus, items from the six primary conditions made up half of the items, and the overall ratio of acceptable to unacceptable sentences was 1:1. Furthermore, since the violations occurred either early in the target sentence (primary conditions) or later in the sentence (control sub-experiments), participants needed to pay attention to the entire sentence in order to accurately judge the well-formedness of the sentence.
2.3 Procedure

Participants were comfortably seated in a dimly lit testing room about 100 cm in front of a computer monitor. Each two-sentence pair (context sentence plus target sentence) was preceded by a fixation cross. All items appeared in black font on a white screen. Participants pressed a button to initiate presentation of the trial. The context sentence was presented in two self-paced sections in 25 pt font: after reading the first half of the sentence, participants pressed a button to see the second half, which began 180 ms after the button press. After reading the second half of the context sentence, participants pressed a button and the target sentence began 180 ms later. Target sentences were presented one word at a time in 30 pt font. Each word appeared in the center of the screen for 300 ms, followed by 200 ms of blank screen. The final word of each sentence was marked with a period, and 1000 ms later a question mark prompt appeared on the screen. Participants were instructed to read the sentences carefully without blinking and to indicate with a button press whether the sentence was an acceptable sentence of Spanish. Feedback was provided for incorrect responses. Each experimental session was preceded by a 3 trial practice session that included both acceptable and unacceptable sentences. Participants received feedback and were invited to ask clarification questions about the task. The experimental session was divided into 7-8 blocks lasting 15 minutes each, but participants could request additional breaks at their discretion.

2.4 EEG recording

EEG was recorded from 28 Ag/AgCl electrodes, mounted in an electrode cap (Electrocap International): midline: Fz, FCz, Cz, CPz, Pz, Oz; lateral: F3/4, F7/8, FC3/4, FT7/8, C3/4, T7/8, CP3/4, TP7/8, P4/5, P7/8, O1/2. Recordings were referenced to the left mastoid. Additional electrodes were placed on the left and right outer canthus, and above and below the left eye to monitor eye movements. The EEG and EOG recordings were amplified by a SynAmps™ Model 5083 EEG amplifier, and sampled at 1 kHz using an analog bandpass filter of 0.1-70 Hz. Impedances were kept below 5 kΩ.
2.5 EEG analysis

All comparisons were based upon single word epochs, consisting of the 100 ms preceding and the 1000 ms following the start of the presentation of the critical words. Epochs with ocular and other large artifacts were rejected from analysis based on visual screening prior to any further analyses. In addition, epochs from trials in which the participant responded inaccurately were excluded. Five participants’ data were excluded because of accuracy below 75%. Epochs that showed a divergence of more than 75\(\mu\)V from zero were excluded automatically. Data from one participant were excluded because only 22% of critical epochs remained after all these measures were taken. For the remaining 24 participants, after all the exclusion criteria were applied, 65% of the trials were included in the final analysis. This rate is relatively low compared to a typical inclusion rate of about 80% in ERP sentence processing studies. There are two likely reasons for this. First, most ERP studies do not exclude trials to which the participants respond inaccurately on the behavioral task. Second, the fact that participants read context sentences prior to the target sentences likely resulted in a higher rate of ocular artifacts during the target sentences. Once these features of the current study are taken into consideration, the data inclusion rates are consistent with those from other ERP studies of sentence processing.

The waveforms were normalized using a 100 ms pre-stimulus baseline. Averaged waveforms were filtered offline using a 10 Hz low-pass filter for presentation purposes, but all statistical analyses are based on unfiltered data. The following latency intervals were chosen for analysis, based on the intervals used in previous literature and on visual inspection: 300-500 ms (N400), 600-1000 ms (P600). The 0-200 ms interval was also analyzed to test for possible early differences.

For statistical analyses, six regions of interest (ROIs) were used in the ANOVAs, consisting of groups of three electrodes at each ROI: left anterior (F3, FC3, C3), anterior midline (FZ, FCZ, CZ), right anterior (F4, FC4, C4), left posterior (CP3, P3, O1), posterior midline (CPZ, PZ, OZ), right posterior (CP4, P4, O2). These ROIs were organized into the two topographic factors *laterality* (left, midline, right), and *posteriority* (anterior, posterior).
ANOVAs were performed separately for the conditions with *fue* and for those with *estaba*, due to the fact that the length difference between the two auxiliaries could introduce irrelevant differences into the waveforms. This separation of conditions in the analyses is consistent with the experimental hypotheses, which focused on the relation between the semantic-attraction and no-attraction conditions and the control condition within each level of the auxiliary factor.

ANOVAs were performed hierarchically using the within-subjects factor *condition* (*control, semantic-attraction, no-attraction*). All *p*-values reported below reflect the application of the Greenhouse-Geisser correction where appropriate to control for violations of the sphericity assumption (Greenhouse & Geisser, 1959), together with the original degrees of freedom. Due to the large number of possible interactions in this design we discuss only those interactions for which follow-up analyses yielded significant contrasts within the levels of the interacting factors.

3. Results

3.1 Accuracy

Overall accuracy on the behavioral acceptability judgment task for the six primary conditions was 84%. The accuracy on the two acceptable control conditions was higher (*fue*-control: 92%; *estaba*-control: 90%) than for the four anomalous conditions (*fue*+semantic-attraction: 75%; *fue*+no-attraction: 84%; *estaba*+semantic-attraction: 82%, *estaba*+no-attraction: 84%). This difference was likely due to the fact that the anomaly resulted from the addition of a single character in the verb suffix (e.g., passive -*ado* became progressive -*ando*) and therefore may easily have been missed on some trials, especially in trials when the verb was long and the suffix far from the visual fixation point. As noted above, only trials with correct responses were included in the analysis.

The semantic control sub-experiment showed somewhat lower overall accuracy (77%), with average accuracy of 84% in the felicitous condition and only 70% in the infelicitous condition, suggesting that some noun-adjective combinations may have been perceived as more
felicitous than other anomalous items in the study. The syntactic control experiment showed an overall accuracy of 87% (grammatical: 89%; ungrammatical: 85%).

3.2 Control sub-experiments

The results for the semantic and syntactic control experiments conformed to expectations. There was an N400 effect at the adjective in the infelicitous noun-adjective combinations of the semantic control sub-experiment, and a P600 at the nouns with mismatching number agreement in the syntactic control sub-experiment.

3.2.1 Semantic Control Sub-Experiment

In the 0-200 ms interval following the critical adjective in the semantic control sub-experiment there were no differences between conditions. In the N400 interval (300-500 ms) voltages across the entire scalp were more negative in the semantic anomaly condition. This negativity led to a main effect of condition in the overall ANOVA, in addition to significant effects in all ROIs except for the right anterior region, where the difference was only marginally significant. Figure 1 shows an electrode array and F-values are shown in Table 3.

![Figure 1](image.png)

**Figure 1.** Grand average responses at the critical adjective in the semantic control conditions, showing the felicitous condition (blue) and the infelicitous condition (red).
Table 3. ANOVA $F$-values for comparisons of the semantic sub-experiment conditions at the critical adjective.

<table>
<thead>
<tr>
<th>Semantic Felicity</th>
<th>0-200 ms</th>
<th>300-500 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>cond (1,23)</td>
<td>-</td>
<td>8.00 *</td>
</tr>
<tr>
<td>cond × lat (2,46)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cond × post (1,23)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cond × lat × post (2,46)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Effect of condition at individual ROIs:
- Anterior left (1,23): 9.35 **
- Anterior midline (1,23): 7.26 *
- Anterior right (1,23): 3.59 †
- Posterior left (1,23): 9.26 **
- Posterior midline (1,23): 8.26 **
- Posterior right (1,23): 6.25 *

** $p < .01$, * $0.01 < p < .05$, † $0.05 < p < .1$
Numbers in parentheses indicate degrees of freedom.

3.2.2 Syntactic Control Sub-Experiment

Visual inspection suggested that the ERP waveforms diverged already in the 0-200 ms interval, particularly at central and posterior electrodes. This was confirmed by the ANOVA, which indicated a significant effect of condition ($F(1,23) = 5.24, p < .05$) and a marginally significant interaction of condition and posteriority ($F(1,23) = 3.52, p < .1$), due to a positivity in the ungrammatical condition. This difference was marginally significant at the anterior midline and anterior right ROIs (anterior midline: $F(1,23) = 3.26, p < .1$; anterior right: $F(1,23) = 3.98, p < .1$) and was significant at all posterior ROIs (posterior left: $F(1,23) = 6.46, p < .05$; posterior midline: $F(1,23) = 6.43, p < .05$; posterior right: $F(1,23) = 5.99, p < .05$). However, this early difference did not persist to the 300-500ms interval, where no differences were observed at any ROI. Figure 2 shows an electrode array and $F$-values are shown in Table 4.

In the P600 interval (600-1000 ms) ERPs in the ungrammatical condition were more positive than the control condition. The effect was present across the entire scalp, but was largest at posterior and midline electrode sites, resulting in interactions between condition and posteriority and between condition and laterality, and a marginally significant three-way interaction between condition, posteriority and laterality. The effect was significant in all regions of interest.
Although the difference in the early interval was unexpected, it does not undermine the interpretation of the P600 effect, since there was no significant difference in the interval directly preceding the P600 interval, and the mean amplitude of the early difference (~1 µV) was much smaller than the mean amplitude of the P600 effect (~5 µV), averaged across the entire scalp. It is therefore unlikely that the difference in the P600 interval could be the result of pre-existing differences.

**Figure 2.** Grand average responses at the critical noun in the syntactic control conditions, showing the grammatical agreement condition (blue) and the ungrammatical agreement condition (red).

**Table 4.** ANOVA F-values for comparisons of the syntactic sub-experiment conditions at the critical noun.

<table>
<thead>
<tr>
<th>Syntactic Agreement</th>
<th>0-200 ms</th>
<th>300-500 ms</th>
<th>600-1000 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>cond (1,23)</td>
<td>5.24 *</td>
<td>-</td>
<td>46.19 **</td>
</tr>
<tr>
<td>cond × lat (2,46)</td>
<td>-</td>
<td>8.77 **</td>
<td>9.58 *</td>
</tr>
<tr>
<td>cond × post (1,23)</td>
<td>3.52 †</td>
<td>-</td>
<td>12.22 *</td>
</tr>
<tr>
<td>cond × lat × post (2,46)</td>
<td>-</td>
<td>3.72 *</td>
<td>3.43 †</td>
</tr>
</tbody>
</table>

**Effect of condition at individual ROIs:**

Anterior – left (1,23) 3.28 †  -  31.90 **
Anterior – midline (1,23) 3.98 †  -  39.03 **
Anterior – right (1,23) 6.46 *  -  34.48 **
Posterior – left (1,23) 6.43 *  -  42.38 **
Posterior – midline (1,23) 5.99 *  -  47.07 **
Posterior – right (1,23) 5.99 *  -  43.63 **

**p < .01, .01 < p < .05, .05 < p < .1**

Numbers in parentheses indicate degrees of freedom.
3.3 Attraction Conditions with the Auxiliary *Fue*

The auxiliaries *fue* and *estaba* are of different lengths, which could introduce differences that are unrelated to the primary manipulation of the study. For this reason all analyses treated the *fue* and the *estaba* conditions separately.

3.3.1 Subject Noun and Auxiliary

In the analysis of the three *fue* conditions (grammatical control, semantic-attraction and no-attraction) there were no significant effects of condition, or interactions with condition, in either the 0-200 ms or 300-500 ms intervals following presentation of the subject noun or in the corresponding intervals following presentation of the auxiliary.

3.3.2 Critical Verb

The critical verb following the auxiliary *fue* elicited a P600 in both anomalous conditions relative to the acceptable control condition, with no differences in the N400 interval. An ANOVA that compared all three conditions showed no differences in the 0-200 ms interval or in the 300-500 ms interval. In the 600-1000 ms interval both anomalous conditions (semantic-attraction and no-attraction) showed a positivity relative to the control condition. The effect was distributed across the whole scalp and was largest in midline and posterior regions. Figure 3 shows the grand-average waveforms. This pattern of results was reflected in a main effect of condition ($F(2,46) = 3.84, \ p < .05$). To determine the source of this effect, planned pairwise comparisons were made between all possible pairs of conditions (Table 5). This confirmed the visual observation that the semantic-attraction condition and the no-attraction condition were both more positive than the control condition, but did not differ from each other.
Figure 3. Grand average ERP responses at the critical verb in the *fue* conditions, showing the grammatical control (blue), semantic-attraction (black) and no-attraction (red) conditions.

Table 5. ANOVA *F*-values for comparisons of the *fue* conditions at the critical verb.

<table>
<thead>
<tr>
<th><em>fue</em></th>
<th>0-200 ms</th>
<th>300-500 ms</th>
<th>600-1000 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control vs. Semantic-attraction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cond (1,23)</td>
<td>-</td>
<td>-</td>
<td>4.41 *</td>
</tr>
<tr>
<td>cond × lat (2,46)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cond × post (1,23)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cond × lat × post (2,46)</td>
<td>-</td>
<td>-</td>
<td>2.85 †</td>
</tr>
<tr>
<td><strong>Effect of condition at individual ROIs:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior left (1,23)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anterior midline (1,23)</td>
<td>-</td>
<td>-</td>
<td>4.87 *</td>
</tr>
<tr>
<td>Anterior right (1,23)</td>
<td>-</td>
<td>-</td>
<td>3.88 †</td>
</tr>
<tr>
<td>Posterior left (1,23)</td>
<td>-</td>
<td>-</td>
<td>4.13 *</td>
</tr>
<tr>
<td>Posterior midline (1,23)</td>
<td>-</td>
<td>-</td>
<td>5.15 *</td>
</tr>
<tr>
<td>Posterior right (1,23)</td>
<td>-</td>
<td>-</td>
<td>3.36 †</td>
</tr>
<tr>
<td><strong>Control vs. No-attraction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cond (1,23)</td>
<td>-</td>
<td>-</td>
<td>7.61 *</td>
</tr>
<tr>
<td>cond × lat (2,46)</td>
<td>-</td>
<td>2.71 †</td>
<td>-</td>
</tr>
<tr>
<td>cond × post (1,23)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cond × lat × post (2,46)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Effect of condition at individual ROIs:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior left (1,23)</td>
<td>-</td>
<td>-</td>
<td>8.26 **</td>
</tr>
<tr>
<td>Anterior midline (1,23)</td>
<td>-</td>
<td>-</td>
<td>8.10 **</td>
</tr>
<tr>
<td>Anterior right (1,23)</td>
<td>-</td>
<td>-</td>
<td>6.12 *</td>
</tr>
<tr>
<td>Posterior left (1,23)</td>
<td>-</td>
<td>-</td>
<td>7.29 *</td>
</tr>
<tr>
<td>Posterior midline (1,23)</td>
<td>-</td>
<td>-</td>
<td>6.12 *</td>
</tr>
<tr>
<td>Posterior right (1,23)</td>
<td>-</td>
<td>-</td>
<td>4.93 *</td>
</tr>
<tr>
<td><strong>Semantic-attraction vs. No-attraction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cond (1,23)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Interaction</td>
<td>0-200 ms</td>
<td>200-500 ms</td>
<td>500-1000 ms</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>cond × lat (2,46)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cond × post (1,23)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cond × lat × post (2,46)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Effect of condition at individual ROIs:**
- Anterior left (1,23) - - -
- Anterior midline (1,23) - - -
- Anterior right (1,23) - - -
- Posterior left (1,23) - - -
- Posterior midline (1,23) - - -
- Posterior right (1,23) - - -

**p < .01, * .01 < p < .05, † .05 < p < .1**
Numbers in parentheses indicate degrees of freedom.

### 3.4 Attraction Conditions with the Auxiliary *Estaba*

Visual inspection suggested that the ERPs in the *estaba* conditions may have diverged prior to the critical verb, and therefore ANOVAs were performed on the 0-200 ms and 300-500 ms intervals after presentation of the subject noun, and on the 0-200 ms, 300-500 ms, and 600-1000 ms intervals after presentation of the auxiliary. The 300-500 ms interval following presentation of the auxiliary corresponds to the interval immediately preceding the presentation of the critical verb.

#### 3.4.1 Subject Noun

At the 0-200 ms interval following presentation of the subject noun the three *estaba* conditions showed no significant main effects of condition or interactions with condition. In the 300-500 ms interval following the subject noun the no-attraction condition showed a positivity relative to the control and semantic-attraction conditions at posterior ROIs. The positivity was reflected in a marginally significant interaction between condition and posteriority ($F(2,46) = 2.69, p < .1$). The control and semantic-attraction conditions were closely matched at this interval, and planned pairwise comparisons confirmed that these conditions did not differ. The comparison of the grammatical control and no-attraction condition showed a marginally significant interaction between condition and posteriority ($F(1,23) = 3.04, p < .1$), although this did not reach significance in any of individual ROIs. The comparison of the semantic-attraction and no-attraction conditions showed a significant interaction between condition and posteriority ($F(1,23) = 5.35, p < .05$), with significant or marginally significant effects at each of the ROIs.
posterior ROIs (posterior left $F(1,23) = 3.88, p < .1$; posterior midline $F(1,23) = 4.12, p < .1$; posterior right $F(1,23) = 4.53, p < .05$).

The contrast between the no-attraction condition and the two other conditions at the subject noun was not predicted. Although the subject nouns differed across conditions within each item set, the nouns were counterbalanced across item sets such that 96% of the subject nouns used in the semantic attraction condition also appeared as the subject noun in a different item set in the no-attraction condition. All subject nouns corresponded to existing discourse referents, as they had already been introduced in the context sentence. It is possible that the subject nouns differed across conditions in terms of how likely they were to appear as subject following the context sentence, and this could be responsible for the difference in the \textit{estaba} no-attraction condition at this word, but that should have led to a similar effect in the \textit{fue} conditions, where no such differences were observed.

3.4.2 Auxiliary

ERPs in the \textit{estaba} conditions showed differences at the auxiliary. An early posterior positivity in the no-attraction condition matched the polarity and scalp distribution of the effect observed immediately prior to the auxiliary, and so is unlikely to reflect an independent response to the auxiliary. In contrast, a broad positivity in both anomalous conditions that started around 300 ms was likely a response to the auxiliary itself. This effect is probably a reflection of the compatibility between the subject noun and the auxiliary: when \textit{estaba} followed an inanimate subject (semantic attraction and no-attraction conditions) it elicited a positivity relative to the condition where it followed an animate subject (control condition). \textit{Estaba} is frequently used in active progressive constructions with animate subjects, and hence the positivity may reflect the low expectancy for \textit{estaba} following an inanimate subject noun. Figure 4 shows grand-average waveforms in all three conditions. Setting aside the early posterior positivity in the no-attraction (red) condition, which was a continuation of a response to the subject noun, the figure shows that the positivity affected the two anomalous conditions similarly. This is most apparent at anterior
electrodes where the ERPs are not contaminated by effects at the previous word, and in the later intervals at posterior electrodes, where the responses to the two anomalous conditions converge. Of course, the interpretation of these results must be treated with some caution, due to the effect at the previous word.

Statistical analyses showed that in the 0-200 ms interval there was a marginally significant effect of condition and a significant interaction between condition and posteriority. Follow-up pairwise comparisons showed that these effects reflected continuations of the differences already seen following the subject noun: the no-attraction condition was more positive than the other two conditions at posterior ROIs, and the control and semantic-attraction conditions did not differ from each other.

In the 300-500 ms interval following the auxiliary there was a three-way difference between the conditions. ERPs in the no-attraction condition continued to be more positive than in the semantic-attraction condition, which in turn were more positive than those in the control condition. These differences were reflected in a significant main effect of condition. This difference between the semantic-attraction and no-attraction conditions was present at posterior sites and not at anterior sites, as reflected in a significant interaction between condition and posteriority. At anterior sites both anomalous conditions elicited a very similar positivity relative to the control condition.

In the 600-1000 ms interval following the auxiliary there was a main effect of condition and no interaction of condition with either topographic factor. This reflected a broad positivity that was present in the semantic-attraction and no-attraction conditions, relative to the control condition. Planned pairwise comparisons showed that the difference between the anomalous conditions and the control condition was reliable at all regions of interest, and that the semantic-attraction condition and the no-attraction condition did not differ from one another at any region.

Table 6 shows $F$-values for the pairwise comparisons of the three conditions.

The positivity at the auxiliary makes it more difficult to interpret any effects that might be observed at the immediately following verb region. However, the effect suggests that
participants were immediately sensitive to the relation between the auxiliary *estaba* and the animacy of the subject noun, and thus it is directly relevant to our questions about the incremental integration of information from open and closed-class words during sentence comprehension. The positivity strongly suggests that the comprehension system does not wait until the verb to evaluate the most likely semantic relations among words in the sentence.

**Figure 4.** Grand average ERP responses at the auxiliary in the *estaba* conditions, showing the grammatical control (blue), semantic-attraction (black) and no-attraction (red) conditions.

**Table 7:** ANOVA *F*-values for comparisons of the *estaba* conditions at the auxiliary

<table>
<thead>
<tr>
<th>Condition</th>
<th>0-200 ms</th>
<th>300-500 ms</th>
<th>600-1000 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estaba (2,46)</td>
<td>3.25†</td>
<td>6.32**</td>
<td>6.49**</td>
</tr>
<tr>
<td>Cond × lat (4,92)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond × post (2,46)</td>
<td>4.18*</td>
<td>6.05**</td>
<td></td>
</tr>
<tr>
<td>Cond × lat × post (4,92)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Control vs. Semantic-attraction**

<table>
<thead>
<tr>
<th>Condition</th>
<th>0-200 ms</th>
<th>300-500 ms</th>
<th>600-1000 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cond (1,23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond × lat (2,46)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond × post (1,23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond × lat × post (2,46)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Effect of condition at individual ROIs:**

<table>
<thead>
<tr>
<th>Region</th>
<th>0-200 ms</th>
<th>300-500 ms</th>
<th>600-1000 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior – left (1,23)</td>
<td></td>
<td>5.20*</td>
<td>20.27**</td>
</tr>
<tr>
<td>Anterior – midline (1,23)</td>
<td></td>
<td>9.44**</td>
<td>22.10**</td>
</tr>
<tr>
<td>Anterior – right (1,23)</td>
<td></td>
<td>7.97*</td>
<td>15.11**</td>
</tr>
<tr>
<td>Posterior – left (1,23)</td>
<td></td>
<td>3.52†</td>
<td>11.73**</td>
</tr>
<tr>
<td>Posterior – midline (1,23)</td>
<td></td>
<td>4.74*</td>
<td>13.98**</td>
</tr>
<tr>
<td>Posterior – right (1,23)</td>
<td></td>
<td>3.97†</td>
<td>10.39**</td>
</tr>
</tbody>
</table>
Control vs. No-attraction

<table>
<thead>
<tr>
<th>Condition</th>
<th>Effect Size</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cond × Lat (2,46)</td>
<td>6.97**</td>
<td>14.25**</td>
</tr>
<tr>
<td>Cond × Post (2,46)</td>
<td>7.67*</td>
<td>5.20*</td>
</tr>
</tbody>
</table>

Effect of condition at individual ROIs:

- Anterior – left (1,23) | 4.79* | 5.43* |
- Anterior – midline (1,23) | 7.98* | 6.20* |
- Anterior – right (1,23) | 7.57* | 5.88* |
- Posterior – left (1,23) | 11.71** | 18.28** | 12.81** |
- Posterior – midline (1,23) | 14.92** | 23.79** | 17.26** |
- Posterior – right (1,23) | 13.98** | 21.06** | 16.00** |

Semantic-attraction vs. No-attraction

<table>
<thead>
<tr>
<th>Condition</th>
<th>Effect Size</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cond × Lat (2,46)</td>
<td>5.71*</td>
<td>12.23**</td>
</tr>
<tr>
<td>Cond × Post (2,46)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Effect of condition at individual ROIs:

- Anterior – left (1,23) | - | - | - |
- Anterior – midline (1,23) | - | - | - |
- Anterior – right (1,23) | - | - | - |
- Posterior – left (1,23) | 4.82* | - | - |
- Posterior – midline (1,23) | 5.55* | 3.97† | - |
- Posterior – right (1,23) | 5.98* | 4.55* | - |

**p < .01, * .01 < p < .05, † .05 < p < .1
Numbers in parentheses indicate degrees of freedom.

3.4.3 Verb

Visual inspection suggested that the semantic-attraction condition may have exhibited a positivity relative to the other two conditions from about 300 ms after the critical verb, but the ANOVAs revealed no significant differences at any interval following the verb (0-200 ms, 300-500 ms, 600-1000 ms). This lack of reliable effects may appear puzzling in light of the fact that the semantic anomalies were not subtle and hence are unlikely to have been overlooked by participants. Additionally, it should be noted that the ERPs for the anomalous conditions are based on the 84% of trials in which participants correctly judged the sentences to be unacceptable. The lack of differences at this region likely is due to a combination of the difference at the preceding auxiliary and the standard baselining procedure that sets all averaged waveforms to be matched in a 100 ms pre-stimulus interval. Figure 5 shows that at the auxiliary there was already a difference between the control condition (animate subject noun) and the two anomalous conditions (inanimate subject nouns), and that this difference was long-lasting and affected both anomalous conditions in the same fashion (once the additional early positivity in
the no-attraction condition elicited by the subject noun is taken into consideration). Thus, the grand-average waveforms shown in Figure 5 do not indicate that there was no processing disruption during the presentation of the verb, but rather that there was no additional disruption beyond that caused by the combination of the subject noun and the auxiliary.

![Figure 5](image-url)

**Figure 5.** Grand average responses at the critical verb in the *estaba* conditions, showing the grammatical control (blue), the semantic-attraction (black) and the no-attraction (red) conditions.

4. Discussion

The aim of this study was to clarify whether findings of thematic P600 effects motivate the addition of an independent, compositional semantic analyzer to the processing architecture, or whether those findings are also compatible with a single analyzer that rapidly integrates diverse sources of information (syntactic, semantic, contextual) as soon as they become available. Previous research (e.g., Friederici & Frisch, 2000; Kim & Osterhout, 2005; Kolk et al., 2003; Kuperberg et al. 2003; Kuperberg et al., 2006; Kuperberg et al., 2007; van Herten et al., 2005; van Herten et al., 2006) has suggested that the processor may be able to use word meanings to propose interpretations that are incompatible with the surface structure of a sentence. However, some of the key evidence for this claim comes from findings that are open to alternative interpretations. Findings in English of P600s elicited by sentences like *The hearty*
meal was devouring … suggest that comprehenders consider a passive interpretation despite contradictory syntactic information (Kim & Osterhout, 2005). This evidence has been taken to show that the language comprehension system incorporates a semantic analyzer that is not directly parasitic upon the syntactic analyzer. However, as described above, what appears to be evidence for an independent semantic analyzer may actually arise from the relative timing with which the various syntactic and semantic pieces of information are perceived and integrated into the ongoing analysis. Furthermore, it has become unclear whether the presence of a plausible thematic relationship among open class words is the key factor in eliciting the thematic P600. The results from Kim and Osterhout (2005) suggested that the P600 is selectively sensitive to the presence of a plausible thematic relationship. However, other studies using different structural configurations (e.g., van Herten et al., 2006, Experiment 2; Kuperberg et al., 2006, 2007) demonstrate that, at least under some circumstances, the P600 may instead be sensitive to specific features of surface syntactic and semantic analyses, such as lexical association or selectional restrictions (e.g., animacy). The current study aimed to determine whether the thematic P600 results provide evidence for an independent semantic analyzer by presenting thematic anomalies in environments that tested (a) whether the language processor identifies plausible thematic relations despite contradictory syntactic information that is available prior to the semantic composition process, and (b) whether the thematic P600 effect is specifically sensitive to the plausibility of thematic relations rather than to lower-level lexical-semantic association effects.

Because relatively few ERP studies have been conducted in Spanish, the experiment also included two control sub-experiments to confirm that syntactic and semantic anomalies elicit canonical N400 and P600 effects. The semantic sub-experiment manipulated the felicity of the semantic relation between a noun and a following adjective, and the syntactic sub-experiment manipulated the correctness of number agreement between a determiner and a noun. The results of the control sub-experiments showed that, as expected, the canonical N400 and P600 effects are also found in Latin American Spanish. An N400 was elicited at the adjective when the
semantic relationship between the noun and the following adjective was infelicitous. Likewise, a P600 was elicited at the noun when the number agreement on the determiner and the noun mismatched. In addition to the P600 effect, there was also an early difference in the 0-200 ms interval following the noun. Although this difference was unexpected, it does not undermine the interpretation of the P600 effect because there was no significant difference in the interval directly preceding the P600 interval and because the mean amplitude of the early difference was much smaller than the P600 effect. It is therefore unlikely that the difference in the P600 interval could be the result of pre-existing differences. These canonical N400 and P600 effects enable us to interpret the presence or absence of similar effects in our primary comparisons in terms of characteristics of the processing architecture, rather than in terms of fundamental differences in electrophysiological responses in Spanish relative to English, Dutch, and German.

In the conditions of primary interest we manipulated the thematic fit between an inanimate subject and a verb. This manipulation follows the logic used by Kim and Osterhout (2005), but it has been tested in relatively few studies. We also introduced a novel manipulation of auxiliary type, using the Spanish past tense auxiliaries *fue* (infinitive: *ser*), which is more frequently used in passive constructions, and *estaba* (infinitive: *estar*), which is more frequently used in active progressive constructions. In addition, our study presented a context sentence prior to each target sentence that introduced the three potential subject nouns for the target sentences. This was motivated by the felicity conditions on the use of passive constructions in Spanish, but it had the additional effect of reducing the effects of lexical availability and global semantic association that may have affected previous studies on the thematic P600.

Results from the conditions with the auxiliary *fue* showed a P600 in both the semantic-attraction and the no-attraction conditions, demonstrating that the thematic P600 is also found in Spanish. The P600 observed in the semantic-attraction condition is consistent with results from other thematic P600 studies (e.g., Kim & Osterhout, 2005; Kuperberg et al., 2006, 2007). This condition was closely comparable to the semantic-attraction condition in Kim and Osterhout’s study. In Kim and Osterhout’s experiment the auxiliary was highly compatible with a passive
structure, and in the current experiment the auxiliary was biased toward a passive structure. The P600 observed in the no-attraction condition with \textit{fue} contrasts with the results of Kim and Osterhout’s study, in which the no-attraction condition did not elicit a statistically reliable P600. However, our finding is consistent with other studies in which a P600 was elicited even in the absence of a plausible thematic relationship among the open class words in the sentence (e.g., Kuperberg et al., 2006, 2007; van Herten et al., 2006).

Thus, the results from the \textit{fue} conditions showed a lack of selectivity of the P600 to the thematic fit between the inanimate subject noun and the verb. This finding is surprising under any account that predicts that the thematic P600 is only modulated by semantic attraction between the noun and the verb. There are at least two possible accounts that might explain this discrepancy.

Firstly, the Spanish auxiliary \textit{fue} provides a stronger cue to the passive than the English auxiliary \textit{was}. In the Spanish case, this may lead the processor to pursue a passive analysis irrespective of the goodness-of-thematic-fit between the noun and the verb. Under this account, when the processor reaches the progressive verbal suffix –\textit{ando}, the structural information carried by the suffix contradicts the processor’s current candidate analysis in both the semantic-attraction and the no-attraction conditions, resulting in a P600 in both conditions. In other words, this account assumes that the non-selectivity of the P600 that we observed in Spanish should not generalize to English or other languages with less strongly biased passive cues.

The second possible explanation for the results from the \textit{fue} conditions is that the P600 is not sensitive to the thematic relationship between the noun and the verb but rather is sensitive to specific properties of surface syntactic and semantic analyses, such as lexical association or selectional restrictions involving features such as animacy. Although this explanation is not consistent with Kim and Osterhout’s interpretations of their results, it is consistent with results from other studies that have shown a P600 in the absence of a plausible thematic fit among the open class words of a sentence in other structural configurations (Kuperberg et al., 2007; van Herten et al., 2006; shown in examples (2) and (3) in the Introduction). Furthermore, our lab has
also found that the P600 is not modulated by thematic fit in two other English studies using the same constructions as the Kim and Osterhout studies, including one study that used identical target sentences (Stroud, 2008). There are several possible accounts for which specific surface features are responsible for the P600, and we tentatively suggest that the critical feature is the animacy mismatch between the subject noun and the verb’s need for an agentive subject.

We should emphasize that both of these possible accounts of the non-selectivity of the P600 involve a sentence comprehension architecture that rapidly integrates multiple sources of information as a sentence unfolds.

The conditions with the auxiliary *fue* also showed no N400 effect in any condition, including the no-attraction condition in which the noun was a poor thematic fit for any of the verb’s arguments. As with the presence of the P600 in this condition, this result is inconsistent with the results from Kim and Osterhout’s study, but the absence of an N400 effect has also been attested in other studies in which there was no plausible thematic fit among open class words (Kuperberg et al., 2007; van Herten et al., 2006). In the current study, the design of the context sentence mitigated differences in semantic association between the subject noun and the verb and overall lexical predictability across conditions. In contrast, in Kim and Osterhout’s materials the semantic-attraction condition generally involved a high level of semantic association between the noun and the verb and the no-attraction condition generally involved a low level of semantic association between the two. This suggests that the N400 effect in that study may have reflected differences in lexical association across conditions rather than the plausibility of the thematic relationship among the open class words. This interpretation is supported by the results of van Herten and colleagues, shown in example (2), which showed an N400 only when the object noun and the verb had a low level of lexical association (van Herten et al., 2006). It is similarly supported by other findings from our lab that show a reduced N400 amplitude when any nearby noun is semantically associated with the verb, regardless of whether it occupies one of the verb’s argument positions (Stroud, 2008). This interpretation of these results is more consistent with the
lexical view of the N400 (Federmeier, 2007; Kutas & Federmeier, 2000; Lau et al., 2008) than with the integration view, as discussed in the Introduction.

Another set of conditions tested for effects of thematic fit in sentences with the auxiliary *estaba*, which is most frequently used in active progressive constructions. In this set of conditions the ERPs at the lexical verb showed no reliable differences between conditions. This lack of differences is initially puzzling, as the semantic-attraction and no-attraction conditions are clearly anomalous and the results from the judgment task showed that participants detected the anomalies with a high degree of reliability (84%). Furthermore, our analyses included only those trials on which the judgment task was answered correctly. Results from the *fue* conditions and from the syntactic and semantic control sub-experiments show that our participants were clearly able to generate familiar N400 and P600 effects within the same experimental session.

However, examination of ERPs at earlier word positions suggests a reason why no effects were observed at the lexical verb. The impact of the inanimate subject nouns and the active-biased auxiliary was seen already at the auxiliary itself, where the anomalous conditions elicited a positivity relative to the control condition, which included an animate subject noun. This positivity had an onset of around 300 ms and was broadly distributed across the scalp.

The exact time-course and distribution of the ERPs elicited by the auxiliary *estaba* is compromised by an additional difference in the no-attraction condition that began 300-500 ms after the subject noun. This posterior positivity overlapped with responses to the auxiliary, making it more difficult to interpret early effects at posterior electrodes following the auxiliary. The additional response to the subject noun in the no-attraction condition was unexpected, since the nouns used in the semantic-attraction and no-attraction conditions were almost completely counterbalanced across item sets, and since all potential subject nouns were introduced in the context sentence and thus corresponded to existing discourse referents. It is possible that the additional response reflected differences in discourse status of the possible subject nouns in the context sentence, but this must remain a speculation at present.
Notwithstanding the concern raised by the overlapping effect from the subject noun in the no-attraction condition, the responses at the auxiliary *estaba* suggest that the semantic-attraction condition and the no-attraction condition affected the ERPs in exactly the same fashion. At anterior electrodes, where there was no confound due to effects from the previous word, the positivity in the two anomalous conditions began around 300 ms after the auxiliary and was closely matched in the two conditions. At posterior electrodes the responses in the two anomalous conditions diverged in the 0-200 ms and 300-500 ms intervals, probably due to the effect at the previous word, but at the 600-1000 ms interval the anomalous conditions showed a similar positivity relative to the control condition. Overall, then, this suggests that the ERPs showed effects of animacy, but no effect of the thematic fit between the subject noun and the verb, just as in the conditions with the auxiliary *fue*. This conclusion must, of course, be treated with some caution, due to the effects of overlapping responses to successive words.

The effect of subject animacy observed at the auxiliary *estaba* suggests that different sources of syntactic and semantic information are rapidly integrated on-line. The fact that this effect was broad and long-lasting may account for why we observed no additional effects of the anomaly at the lexical verb. It is unclear whether this positivity should be treated as functionally equivalent to the positivity elicited at the lexical verb in the *fue* conditions. The early onset (around 300 ms) and the broad scalp distribution may suggest that this effect is different from standard P600 effects, which tend to exhibit a posterior focus. However, comparison with the ERPs in the *fue* conditions and the syntactic control sub-experiment reveals that the positivities were broadly distributed across the scalp in each case. The positivity at *estaba* certainly has an earlier onset than the positivity in the *fue* conditions, but it is already known that lexical and syntactic factors can affect the timing of late positivities within studies (Gouvea, Phillips, Kazanina, & Poeppel, 2010; Osterhout, Holcomb, & Swinney, 1994; Phillips, Kazanina, & Abada, 2005), and the earlier effect in the *estaba* condition may reflect the high lexical frequency of the auxiliary.
The effect of animacy that we observed at the auxiliary in the \textit{estaba} conditions contrasts with results from its closest English counterpart. Kim and Osterhout’s (2005) first experiment included two control conditions: a passive control (\textit{the hearty meal was devoured…}) and an active control (\textit{the hungry boy was devouring …}). In that study the choice of control did not affect the status of the P600 elicited at the lexical verb in the anomalous condition. Visual inspection suggests that responses were slightly more positive in the passive control condition (inanimate subject) than the active control condition (animate subject), but this difference was not statistically reliable. Consequently, Kim and Osterhout’s subsequent study included only the passive control condition. The large number of conditions in our study made it impossible to include both active and passive control conditions, and we therefore chose to use an active control for the \textit{estaba} conditions because it was easier to make these sentences felicitous following a context sentence that also needed to be compatible with the other five versions of each item. The effect of animacy at \textit{estaba} in our study may reflect the stronger biases associated with Spanish \textit{estar} than with English \textit{be}. \textit{Estar} is commonly used in active progressive constructions, which typically include an agentive subject. It is also used with stage-level predicates (Carlson, 1977), which indicate temporary states – e.g., \textit{I am happy right now}, or \textit{I am sick}, whereas the auxiliary \textit{ser} is used with individual-level predicates that indicate more permanent states, as in \textit{I am a happy person in general}, or \textit{I am tall}. The fact that \textit{estar} is used with stage-level predicates may contribute to the expectation for its subject to be animate. Additionally, the imperfective verbal suffix –\textit{aba} used in \textit{estaba} may also have contributed to an expectation for an animate subject, since \textit{estaba} is used with the progressive voice, and inanimate nouns may occur less frequently with the progressive morphology because they are poor agents.

Taken together, the results from both the \textit{fue} and \textit{estaba} conditions in the current study show effects of mismatches between subject animacy, auxiliary bias, and verb voice, and these effects occur as soon as these effects are detectable in the surface form of the sentence. Meanwhile, we find no clear effects of thematic attraction between the subject and the verb.
These results can be explained in terms of a language processing architecture that rapidly integrates information from different sources, including the surface syntax of a sentence, with no need to invoke an independent semantic analyzer that ignores surface syntactic cues. Such architectures are entirely standard in psycholinguistics and neurolinguistics (e.g., Ferreira & Clifton, 1986; Friederici, 2002; Hagoort, 2008; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994).

We should emphasize that the current study does not account for all P600 effects that have been attributed to an independent semantic processing stream. Our discussion here has focused on evidence from sentences with verb-argument animacy mismatches, which have made up a substantial proportion of previous studies on the thematic P600 in English and Dutch. Our current findings do not extend to cases of subject-object-verb (SOV) sentences with reversed animate arguments, such as the Dutch counterpart of The fox that at the poachers hunted (van Herten et al, 2005). Such reversals elicited a P600 in Dutch, despite their syntactic well-formedness, and our lab has found similar effects in SOV sentences in Mandarin Chinese (Chow & Phillips, in prep.). Such cases await further investigation.

5. Conclusion

The current study tested whether the language processor identifies plausible thematic relations despite contradictory syntactic information that is available prior to the semantic composition process, and whether the ‘thematic P600’ effect is specifically sensitive to the plausibility of thematic relations, rather than to lower-level lexical-semantic association effects. ERP recordings from Spanish showed that a broadly distributed late positivity is elicited as soon as comprehenders detect an unlikely combination of the animacy of a subject noun, the bias of an auxiliary, or the voice morphology (passive vs. active) of a lexical verb. In conditions with the passive-biased auxiliary fue the positivity was elicited at the lexical verb, and in conditions with the active-biased auxiliary estaba the positivity was elicited at the auxiliary itself. In an important previous study by Kim and Osterhout, manipulation of thematic attraction between a
subject and a verb elicited contrasting ERP responses, and the contrast has been interpreted as a key piece of evidence for an independent semantic analyzer that considers meanings that are not licensed by the syntactic structure of a sentence. In the current study we found no effects of the manipulation of thematic attraction, and thus no evidence that the language processor considers interpretations that are inconsistent with the surface structure of the sentence. We suggested that the effect of thematic attraction found in previous studies may be due to lower-level lexical association effects. Further we suggested that the P600 effect is sensitive to features of surface syntactic and semantic analyses, such as selectional restrictions involving animacy. In sum, this study did not find evidence that the language processor considers interpretations that are not licensed by the syntactic structure of the sentence. Also, we found that once differences in lower-level lexical-semantic association were minimized, the thematic anomalies did not elicit an N400 response, regardless of whether there was a plausible thematic relationship between the noun and the verb. We propose that rather than motivating an independent, semantic compositional analyzer, these results reflect a language processor that rapidly integrates information from multiple sources (e.g., syntax, semantics, discourse, lexical probabilities) to continuously update its interpretation of an incoming sentence.

**Acknowledgments**

Preparation of this paper was supported in part by fellowships to CS from the Social Sciences and Humanities Research Council of Canada and the Alberta Government, and by grants to CP from the National Science Foundation (BCS-0948554, DGE-0801465). We would like to thank Iván Ortega-Santos, Leticia Pablos, Angel Gallego, Laura Fuentes, Antonia Carrasco, and Juan Uriagereka for help in creation of the Spanish experimental materials, and Gina Kuperberg, Norbert Hornstein, Albert Kim, Herman Kolk, David Poeppel, and Wing Yee Chow for valuable discussion of earlier versions of this work.

**References**


