Available online at www.sciencedirect.com
ScienceDirect
Biological Psychology 75 (2007) 8–18
BIOLOGICAL PSYCHOLOGY
www.elsevier.com/locate/biopsycho

The interplay of heuristics and parsing routines in sentence comprehension: Evidence from ERPs and reaction times

Constance Th.W.M. Vissers *, Dorothee J. Chwilla, Herman H.J. Kolk

Institute for Cognition and Information, Radboud University, Nijmegen, The Netherlands

Received 6 December 2005; accepted 10 October 2006
Available online 27 November 2006

Abstract

Semantic anomalies like ''the fox that hunted the poacher'' elicit P600 effects. Kolk et al. [Kolk, H.J., Chwilla, D.J., Van Herten, M., Oor, P.J.W., 2003. Structure and limited capacity in verbal working memory: a study with event related potentials. Brain and language, 85(1), 1–36] proposed that this P600 effect is triggered by a conflict between the outcome of a lexical strategy with that of the parsing routine. Specifically, when the lexical strategy indicates that the poacher hunted the fox, the full parse leads to the conclusion that the fox was the one who did the hunting. We tested this hypothesis by replicating the study cited above but manipulating the context by means of instruction. Participants were informed that semantic anomalies were created on purpose and that they should not be misled by these anomalies but instead focus on syntax or sentence structure. This instruction led to a strong reduction in P600 effect. This result supports the view that expectations play an important role in the generation of P600 effects to semantic anomalies, as proposed by Kolk et al. [Kolk, H.J., Chwilla, D.J., Van Herten, M., Oor, P.J.W., 2003. Structure and limited capacity in verbal working memory: a study with event related potentials. Brain and language, 85(1), 1–36].

Keywords: Language; Sentence processing; Event-related potentials; Parsing routines; Heuristics; P600; N400

1. Introduction

Under normal circumstances we rarely experience difficulties in understanding sentences in our mother tongue. People comprehend sentences rapidly and without conscious effort. However, numerous problems have to be resolved during the process of understanding a sentence. Sentences are not just strings of words, linked together in a random way. Word by word, information from several sources becomes available that readers and listeners must use, specifically: pragmatic, semantic and syntactic information. Theories about the way in which these sorts of information are represented, retrieved and combined during language processing have urged different lines of psycholinguistic research during the last decade. In the 1960s, most psycholinguistics shared Chomsky's view that syntactic processing proceeds independently of semantic information. This view is exemplified by Chomsky's example sentences (Chomsky, 1957, p.15): 'Colorless green ideas sleep furiously.' and 'Furiously sleep ideas green colorless.' These two sentences are equally absurd, but one will recognize that only the former is syntactically correct. Chomsky (1957, p.15) says: 'The notion ''grammatical'' cannot be identified with ''meaningful'' or ''significant'' in any semantic sense.' Fodor (1983) assimilated the concept of independent syntactic processing into the general idea that syntactic processing is modular. Modular processes are algorithmic, autonomous and operate bottom-up. These processes are not guided by world knowledge, beliefs or expectations. So, according to Fodor, syntactic processes are not guided by the meaning or plausibility of the sentence.

The so called syntax-first theories embody the above-mentioned claim that the syntactic module is restricted to the domain of syntactic information and is immune to non-syntactic information. In this module, an initial commitment is made to a single syntactic structure of the sentence as it has developed so far, on the basis of syntactic information alone. Semantic and pragmatic information is represented in another module and is only activated at a later processing stage (e.g., Frazier and Fodor, 1978; Ferreira and Clifton, 1986).

Constraint satisfaction models have been proposed as a competitor of syntax-first models. According to these models,

* Corresponding author at: NICI, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands. Tel.: +31 243612652; fax: +31 243616066.
E-mail address: c.vissers@nici.ru.nl (C.T.W.M. Vissers).

0301-0511/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.biopsycho.2006.10.004

information from all sources interacts continuously during the formation of an internal interpretation. That is, all kinds of information provided by the incoming words (e.g., context, discourse and semantic information) can jointly affect the activation of different syntactic alternatives (e.g., McDonald et al., 1994; St. John and McClelland, 1990; Trueswell and Tanenhaus, 1994).

Ever since Fodor (1983) argued that parsing is basically a reflex, most models of sentence comprehension have assumed that an interpretation must be based on an initial syntactic structure, even if the activation level of that structure can be influenced by nonsyntactic sources of information. This implies that, according to both syntax-first models and constraint satisfaction models, every interpretation is complete, detailed and accurate and can not be based on shallow, inaccurate or incomplete processing. However, the meaning people derive from a sentence is often not a reflection of its true content. Hence, what current models are missing is an architectural component that can explain cases in which people use strategies or engage in heuristic processing of sentences that, may then result in an inaccurate interpretation (Ferreira et al., 2002; Ferreira, 2003).[1]

Researchers in other domains of cognition, such as decision making and reasoning, have argued that human behavior is at least partially driven by heuristic processing (Gigerenzer and Goldstein, 1996; Gigerenzer et al., 1999). Gigerenzer and his colleagues have proposed the idea of bounded rationality. They argue that humans often find themselves in situations which force them to make inferences about the world under limited time, knowledge, and computational power. Models of rational inference do not take these limitations into account and are therefore unrealistic. They show that simple heuristics can match or even outperform classical models of rational inference. These simple heuristics are frugal because they exploit the structure of environments and only require the use of a small proportion of the available information. As a consequence, they are thought to be fast because information search is less computationally demanding.

How might heuristics be used during language processing? One proposed heuristic is referred to by Bever (1970) as 'strategy C', and by Ferreira et al. (2002) as 'the plausibility strategy' and will be referred to below as 'the lexical strategy'. This strategy is a semantic heuristic which states that readers depend heavily on their knowledge of the meaning of individual content words, which provide a strong basis for the most plausible interpretation. Hence, according to this strategy, readers depend on schemas in long-term memory or world knowledge. Evidence for the existence of such a bias comes from work with aphasic patients (Saffran et al., 1989). In particular, these patients have great difficulty rejecting sentences like *the painting disliked the artist*, even though this is a simple active construction, a sentence type they have little difficulty with in other tests. Another heuristic is the canonical 'word order strategy' (Townsend and Bever, 2001; Bever et al., 1998). This heuristic says that a noun phrase preceding a verb is taken as the subject of that verb. Similarly, a noun phrase following a verb is taken to be the object of the verb. Since this order is present in the majority of sentences in the English language, this is a very advantageous strategy.

Using a plausibility strategy implies a process of integration: one takes a set of content words and attempts to fit these words into a meaningful whole. Now, precisely such a process of meaning integration is what appears to set off a well-known language-relevant ERP component: the N400. The N400 is a negative voltage peak that reaches its maximum amplitude around 400 ms after the onset of open class words (for reviews: Kutas and Van Petten, 1994; Kutas and Federmeier, 2000; Kutas and Schmitt, 2003). Pairs of words that do not fit well together semantically – e.g., 'cat' and 'rose' – elicit larger N400 amplitudes than pairs of words that do – e.g., 'cat' and 'dog' – (e.g., Chwilla et al., 1995, 1998). That this effect truly reflects meaning integration was further confirmed by the results of a study of Chwilla et al. (2000). These authors found that indirect semantic relationships between words (e.g., Prime: 'lion', Target: 'stripes', Mediator [not presented]: 'tiger') elicited an N400 effect only if these indirect relationships were the strongest relationships in a list. In contrast, if the list also contained directly related pairs (e.g., 'girl' and 'boy'), no effect of indirect relationships was found. This led the authors to conclude that the participants attempted to find semantic coherence in a pair of words at the highest level of coherence that the material permits.

The plausibility strategy described above refers to a process of meaning integration between a set of content words that occurs in a sentence. Therefore, the question arises whether the process of meaning integration for pairs of isolated words that is reflected by the N400 is similar to the integration process also reflected by the N400 within sentences. This question has been investigated by Kutas (1993) by comparing the N400 effects to word pairs and sentences in the same group of subjects. The outcome was that the N400 relatedness effect does not depend upon whether the words occur in a list or in a sentence: as the latency, amplitude and overall shape of the N400 effects in the two conditions were very similar. It therefore seems likely that a process of meaning integration as we described above, in which participants attempt to find maximal coherence between content words, also takes place within sentences. This is exactly what the plausibility strategy embodies.

The evidence reviewed above suggests that a comprehensive theory of language comprehension should assume that simple processing heuristics are used in addition to syntactic algorithms. The question remains as to how these heuristics and parsing systems are coordinated. According to the Late Assignment of Syntax Theory (LAST) (Townsend and Bever, 2001; Bever et al., 1998), the initial semantic analysis of sentences proceeds on the basis of statistically sensitive perceptual strategies or heuristics (the authors only mention the word-order strategy here, but the same reasoning could be

[1] With heuristics we mean that in some situations language users do not take all relevant information into account, syntactic as well as semantic information, but only a specific part. As we will see below, this part may either be word order or the meaning of the set of lexical items.

applied to the plausibility strategy). Townsend and Bever (2001) refer to this analysis as a ''pseudo-parse'', because the comprehension mechanism does not operate as the mere mechanical application of syntactic categories and frames from left to right. Instead, the pseudo-parse is a probabilistic analysis of meaning and form that proposes a likely candidate meaning or conceptual structure. So, the pseudo-parser uses heuristics to create a preliminary hypothesis or 'best-guess' about the input. In addition, a more time-consuming algorithmic analysis proposes a candidate real syntax. This true parser uses the preliminary hypothesis to constrain its initial hypothesized search space. If this algorithmic parser has time to finish all its computations, it will output the complete and correct interpretation of the sentence. Townsend and Bever (2001) state: 'semantics propose, syntax disposes', (p. 271). Ferreira (2003) also proposes a dual route model and argues that the language comprehension system uses a combination of heuristics and syntactic algorithms.

In most sentences, heuristics and syntactic algorithms will produce the same thematic interpretation. However, in particular sentences, the two routes may produce conflicting results. Such a conflict may underlie the ERP findings, obtained by Kolk et al. (2003). They used semantic reversal anomalies which were formed by exchanging the subject and object of semantically acceptable sentences such as (1).

(1) De vos die op de stropers joeg sloop door het bos (original).
The fox that at the poachers hunted[singular] stalked through the woods (literal translation).
The fox that hunted[singular] the poachers stalked through the woods (paraphrase).

It is clear that in these sentences lexical strategies and parsing routines produce different thematic interpretations. Whereas the lexical strategy leads to the interpretation that poachers are hunting foxes, the parsing routines lead to the interpretation that the foxes are hunting the poachers. Although the latter interpretation is not entirely impossible it represents a highly unlikely event based on world knowledge.

Kolk et al. (2003) observed a P600 effect – a late positive potential starting at about 600 ms after the onset of a target word – and not an N400 effect to these semantic reversal anomalies. This is consistent with recent findings from other researchers who, despite differences in sentence material and language (English and Dutch), observed a P600 effect in the absence of an N400 effect to semantically implausible sentences relative to their plausible counterparts (Hoeks et al., 2004; Kim and Osterhout, 2005; Kuperberg et al., 2003, 2006; Van Herten et al., 2005, and see also Vissers et al., 2006). These results seem to challenge the view that P600 effects are reliably elicited by syntactic anomalies whereas N400 effects are elicited by semantic anomalies.[2] How can one account for these unexpected results and how does it relate to the presence of a conflict between algorithmic and heuristic processing routes?

[2] For a discussion of different views on the occurrence of a P600 effect after semantic anomalies, see Van Herten et al. (2005).

Kolk et al. (2003) proposed that the P600 effect reflected the conflict between the thematic interpretations whereas the absence of the N400 effect was due to the fact that the 'lexical' interpretation was the same for plausible and implausible sentences. According to the above described semantic integration view of the N400, the attempt to integrate sets of individual content words into one coherent meaning, is reflected in the N400. Since the word level integration process does not have difficulty integrating the words of the semantic reversals in the study of Kolk et al. (2003) into a coherent meaning (the lexical items in both conditions are the same; the fox and the hunters) readers initially do not notice the anomaly. Hence, no N400 effect was elicited.

Why would a P600 effect follow a conflict between thematic interpretations? Kolk et al. (2003) argued that the language comprehension system attempts to resolve the conflict by reprocessing the sentence to check the memory trace of the input sentence for possible processing errors. In particular, the mismatch between the semantically plausible, highly expected (based on world knowledge) thematic interpretation and the implausible thematic interpretation makes it necessary for the brain to re-attend the unexpected linguistic unit to check upon its veridicality. After all, an inconsistency can have two sources. It can be real, in the sense that an unexpected event has indeed occurred (e.g., man bites dog). On the other hand, it can also stem from a processing error. To prevent integration of erroneous information into the current discourse, the reader will generally check upon the correctness of his or her analysis in case of a conflict. This explains the occurrence of P600 effects to semantic reversal anomalies.

The levels of processing framework has been used to further determine the processing nature of the N400 component (e.g., Chwilla et al., 1995; Besson et al., 1992; Kutas and Hillyard, 1989) and the P600 component (e.g., Gunter and Friederici, 1999; Gunter et al., 1997; Hahne and Friederici, 1997). According to this framework, different task demands are assumed to result in different levels of processing (Craik and Lockhart, 1972) during word processing. For instance, Gunter and Friederici (1999) used a physical task in the shallow processing condition, in which participants had to judge whether a word in a sentence was printed in upper case (a shallow level of processing). The deep processing condition was a grammatical task in which subjects had to judge sentence-final syntactic errors. This task manipulation was effective in modulating the amplitude of P600; the physical judgment task greatly attenuated or eliminated the N400 and P600 following incorrect verb inflections compared to the grammatical judgment task. This was taken to indicate that the P600 mainly reflects controlled syntactic processing.

In the present paper we used a similar approach to investigate whether the P600 to semantic reversal anomalies is modulated by instruction. Rather than drawing the participants' attention to the physical level as Gunter and Friederici (1999) did, we drew the participants' attention to the syntactic level. More specifically, the aim of the present experiment was to test the hypothesis by Kolk et al. (2003) that the P600 effect is a reflection of the control operation triggered

by a mismatch between the thematic interpretations proposed by the semantic heuristic and the syntactic parser. To this aim, we used the same stimulus materials as in the study of Kolk et al. (2003). However, there was one essential difference in the way participants were instructed in the present experiment. In the Kolk et al. (2003), participants were asked to indicate if the sentence was semantically plausible or not. Implausible was defined as semantically unacceptable. In the present experiment, participants were told that semantic reversals had been created on purpose and that they should not be misled by their knowledge of what normally happens in the world, but pay extra attention to ''who does what to whom'' and to evaluate whether this scenario fits well with their world knowledge or not. Note that this focus-on-syntax instruction creates a context in which semantically odd sentences are expected.

As stated above, a P600 effect is assumed to be triggered by a conflict between the outcome of a lexical strategy with that of the parsing routine. From this, we predicted that if the P600 effect after semantic reversal anomalies reported by Kolk et al. (2003) reflects a control process to check for possible processing errors, the P600 effect in the present experiment should be reduced. Since our focus-on-syntax instruction creates a context wherein anomalous interpretations are expected, the discrepancy between the thematic interpretations proposed by the semantic heuristic and the syntactic parse should be diminished and therefore less readily qualified as a possible processing error. Hence, we predicted that our focus-on-syntax instruction should diminish the inclination to re-attend to a possible processing error which should be reflected in a decrease in error rates and a reduction or even elimination of the P600 effect to semantic reversal anomalies.

2. Method

2.1. Participants

There were 38 participants (mean age = 22 years; age range = 18–30; 29 females). All were native speakers of Dutch, had no reading disabilities, were right-handed and had normal or corrected-to-normal vision. Hand dominance was assessed with an abridged Dutch version of the Edinburgh Inventory (Oldfield, 1971). Sixteen participants reported the presence of left-handedness in their immediate family.

2.2. Materials

The semantic list consisted of 68 Dutch sentences with centrally embedded relative clauses. For each sentence, a subject relative (SR) and an object relative (OR) version, a plausible and an implausible version were created, yielding a total set of 272 sentences (see Table 1). Fourteen (out of 68) sentences in the object relative condition employed adjuncts rather than prepositional complements. The semantically anomalous sentences expressed scenarios conflicting with general world knowledge (e.g., foxes are not very likely to be hunting poachers whereas poachers are likely to hunt foxes). The anomalies resulted from reversing the first and the second noun phrase of semantically acceptable sentences. The two noun phrases could both serve as the agent and the patient of the action expressed by the verb ending the relative clause (e.g., foxes and poachers can hunt as well as be hunted). The anomaly was not evident before the relative clause's verb. This was done to ensure that the detection of the anomaly required deep processing of the relative clause, in that it depended on the successful integration of the verb with both noun phrases. In half of the sentences, the two noun phrases had the same grammatical number and in the other half they had a different number (singular or plural). The four versions of each sentence were counterbalanced across lists. Each list contained 17 SR acceptable sentences, 17 SR semantically anomalous sentences, 17 OR acceptable sentences, and 17 OR semantically anomalous sentences. Sixty-eight filler sentences were added to each list: 17 acceptable right-branching sentences, 17 semantically anomalous right-branching sentences (e.g., *De rechter luisterde naar de beklaagde die opkwam voor zijn advocaat*. Literal translation: The judge listened to the defendant who stood up for his lawyer.), 17 acceptable conjunctions and 17 conjunctions with a semantic reversal anomaly (e.g., *De zeehonden doken in het water en vingen de ijsbeer*. Literal translation: The seals plunged into the water and caught the polar bear.).

2.3. Procedure

Participants were seated in a closed chamber. A response device with three push-buttons was fixed on a small table in front of the participant. Sentences were presented in serial visual presentation mode at the center of a PC monitor. Word duration was 345 ms and the stimulus-onset asynchrony (SOA) was 645 ms. Sentence final words were followed by a full stop. The inter-trial interval was 2 s. Words were presented in black capitals on a white background in a 9 cm × 2 cm window at a viewing distance of approximately 1 m. Each sentence was preceded by a fixation cross (duration 510 ms) followed by a 500 ms blank screen.

Participants were told that semantic anomalies had been created on purpose, by reversing the agent and patient of otherwise normal sentences (fox is hunting poacher instead of poacher is hunting fox). They were asked to attend carefully to the structure of the sentences and evaluate who does what to whom. So, participants had to pay close attention to who was the agent and who was the patient of the sentence and indicate whether this fit well with their world knowledge or not. The experimenter pointed out that it was

Table 1
Examples of the acceptable and unacceptable versions of the sentences separately for the 2 levels of complexity

	Plausible sentence	Implausible sentence
Subject relative	De docent die aan de studenten *lesgaf* kwam het lokaal in	De studenten die aan de docent *lesgaven* kwamen het lokaal in.
Word-by-word translation	The teacher who on the students gave[single] lesson entered the room	The students who on the teacher gave[plural] lesson entered the room.
Paraphrase	The teacher who taught[singular] the students entered the room	The students who taught[plural] the teacher entered the room
Object relative	De studenten aan wie de docent *lesgaf* kwamen het lokaal in	De docent aan wie de studenten *lesgaven* kwam het lokaal binnen
Word-by-word translation	The students to whom the teacher lesson gave[singular] entered the room	The teacher to whom the students gave[plural] lesson entered the room
Paraphrase	The students who were taught[singular] by the teacher entered the room	The teacher who were taught[plural] by the students entered the room

important that the participants should not be misled by their knowledge of what normally happens in the world (that is, normally poachers are hunting foxes and foxes are being hunted) but attend to the structure of the sentences. Incongruent was defined as unlikely based on our knowledge of the world. During a short training the detection of semantic incongruencies was practiced with 10 sentences on paper. The experimenter turned the participants' attention to the reversals in each of the 10 sentences and explained how they were constructed. Participants were subsequently instructed to attentively read each sentence presented on the computer screen and to press a button with the dominant index finger if the sentence was congruent with world knowledge and with the other index finger if it was not. See Table 1 for the example sentences.

The list was split up into 5 blocks; there was a brief pause between blocks and each block was preceded by two filler items. Participants had to postpone their acceptability judgment until presentation of a prompt that occurred 1500 ms after the sentence final word. We used a delayed response task to eliminate effects of motor response preparation on the ERPs of interest. Because eye movements distort the EEG recording, participants were trained to make eye movements, e.g., blinks, only in the period that the prompt was present (stimulus duration was 2295 ms). Prompt offset was followed after 705 ms by a fixation cross indicating the start of the next trial.

2.4. EEG data acquisition and analysis

EEG was recorded with 27 tin electrodes mounted in an elastic electrode cap (Electrocap International). Fig. 1 presents the electrode configuration.

The electrode positions included standard International 10–20 system locations over the left and right hemispheres at the frontal (F3, F4, F7 and F8), midline (Fz, Cz, Pz, Oz), parietal (P3, P4) and temporal (T5, T6) sites. Eight extra electrodes were placed at the frontal (F3A, FZA, F4A, F7A, F8A), midline (OZ) and parietal (P3P, P4P) sites. In addition, eight electrodes were placed at non-standard electrode positions previously found to be sensitive to language manipulations (e.g., Holcomb and Neville, 1990): left and right anteriortemporal sites (LAT and RAT: 50% of the distance between T3/4 and F7/8), left and right temporal sites (LT and RT: 33% of the interaural distance lateral to Cz), left and right temporoparietal (LTP and RTP: Wernicke's area and its right hemisphere homologue: 30% of the interaural distance lateral to a point 13% of the nasion-inion distance posterior to Cz), and left and right occipital sites (OL and OR: 50% of the distance between T5/6 and O1/2). The left mastoid served as reference. Electrode impedance was less than 3 kOhms. The electro-oculogram (EOG) was recorded bipolarly; vertical EOG was recorded by placing an electrode above and below the right eye and the horizontal EOG was recorded via a right to left canthal montage. The signals were amplified (time constant = 8 s, bandpass = .02–30 Hz), and digitized online at 200 Hz. Presentation of stimuli and recording of performance data was accomplished by a Macintosh computer.

EEG and EOG recordings were examined for artifacts and for excessive EOG amplitude (>100 μV) extending from 100 ms before the onset of the critical verb ending the relative clause to 1000 ms following its onset. Averages were aligned to a 100-ms baseline period preceding the critical verb. Based on previous studies using the same or similar materials (Kolk et al., 2003; Van Herten et al., 2005), mean amplitudes were calculated in the time windows of 400–500 ms and 650–850 ms to capture N400 and the P600 effects, respectively. Another reason these windows were chosen was to make a direct comparison between our dataset and the dataset of Kolk et al. (2003) possible. For both time-windows, separately for the midline and the lateral sites, repeated measures MANOVAs were conducted with plausibility (plausible versus implausible) and complexity (SR versus OR) as within subject factors. For the midline sites the additional factor was site (Fza, Fz, Cz, Pz, Oz). To further explore the scalp distribution of the ERP effects for the lateral sites we used a region of interest (ROI: anterior versus posterior) by hemisphere by lateral site (F7a/F3a/F7/F3 versus P3/P3p/T5/OL versus F8a/F4a/F8/F4 versus P4/P4p/T6/OR) design. Relevant interactions with site and plausibility were followed up by post hoc Newman-Keuls' tests to assess the significance of contrasts.

The multivariate approach to repeated measurements was used to avoid problems concerning sphericity (e.g., Dien and Santuzzi, 2004; Vasey and Tayer, 1987).

3. Results

3.1. Reaction time and error data

The reaction time (RT) and error data were entered into separate repeated measures multivariate analyses of variance (MANOVAs) with plausibility (plausible versus implausible), and complexity (SR versus OR) as within subject factors. Mean RT and error percentages are presented in Table 2. The RT analysis revealed main effects of complexity, F (1,36) = 21.32, $p < .001$, and plausibility, F (1,36) = 5.86, $p = .021$. As Table 2 shows, the complexity effect indicated that mean RT for OR sentences (570 ms) was longer than mean RT for SR sentences (517 ms). The plausibility effect indicated that mean RT for plausible sentences (560 ms) was longer than mean RT for implausible sentences (527 ms). There was no complexity by plausibility interaction, $F < 1$.

The error analysis revealed main effects of complexity, F (1,36) = 19.21, $p < .001$, and plausibility, F (1,36) = 8.9, $p = .005$. These effects indicated that participants made more errors on OR (10%) than on SR sentences (5%) and that participants made more errors on plausible (9%) than implausible sentences (6%).[3] No complexity by plausibility interaction was observed, $F < 1$.

Compared to the Kolk et al. (2003) study, participants in the present study appeared to have faster reaction times and lower error percentages. To test the significance of this difference we entered both the RT and error data of the Kolk et al. (2003) and the present study into a separate repeated measures MANOVA, with instruction (semantic plausibility judgment versus focus-on-syntax) as between-subject factor and complexity (SR versus OR) and plausibility (plausible versus implausible) as within-subject factors.

3.1.1. Global analyses

The RT analysis revealed a main effect of Instruction, F (1,76) = 33.79, $p < .001$. Mean RT was faster in the present study (543 ms) than in the Kolk et al. (2003) study (861 ms). No interactions were obtained, Fs < 2. The error analysis indicated that participants in the present study were more accurate than those in the Kolk et al. (2003) study, 8% versus

[3] As expected, participants processed the OR sentences more slowly than the SR sentences, in addition they were less accurate on the OR sentences. This pattern is consistent with the more complex syntactic structure of OR sentences. On the other hand, plausible sentences were responded to more slowly than implausible ones and elicited more errors. This pattern appears unexpected since one might expect plausible sentences to be processed faster and/or more accurately, but this pattern was also observed by Kolk et al. (2003). They explained it in the following way. Because participants had to wait until they had read the last word of the sentence before they could know that the sentence was plausible, they had to postpone their answer until the last word of the sentence was presented. However, participants could know that a sentence was implausible as soon as the anomalous verb was read. It is possible, therefore, that in the case of an implausible sentence, the decision could be made earlier in the sentence and participants could already prepare their response during the sentence. Perhaps, this led to faster mean reaction times and lower error percentages for the implausible sentences.

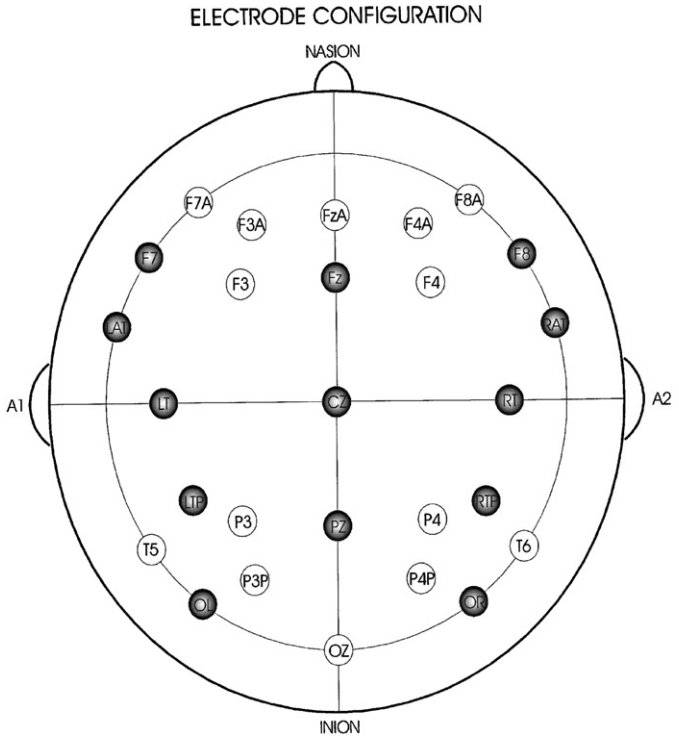

Fig. 1. Electrode configuration used in the present experiment. The electrode configuration of the Kolk et al. (2003) study is displayed by the bold circles.

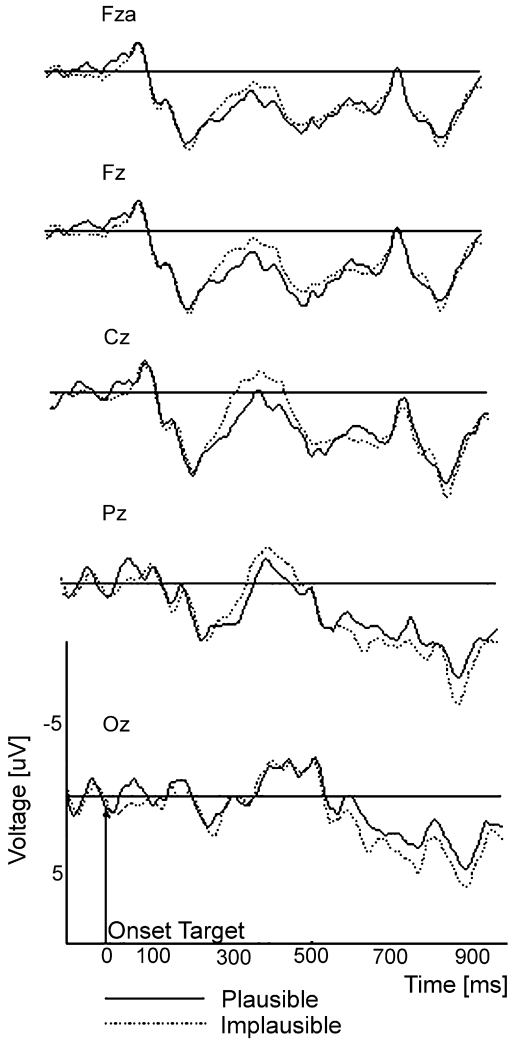

Fig. 2. Grand ERP averages to the critical verb for the midline sites, collapsed over the two levels of complexity. Averages are time locked to the onset of the critical verb, and superimposed for the two levels of Plausibility. Negativity is plotted upwards.

11% of errors respectively, F (1,76) = 5.32, p = .024, indicating the there was no speed accuracy trade off. No further interactions were observed, Fs < 3.5.

The global analyses thus confirmed that the focus-on-syntax instruction affected the behavior of the participants in the predicted direction. Participants were less easily misled by the semantic reversal anomalies when they were previously informed about the presence of semantic reversal anomalies and were instructed to attend to the syntax. This is reflected by the fact that participants in the present study were faster and more accurate than those in the Kolk et al. (2003) study, in which participants' attention was not directed to the sentence structure.

3.2. Event-related potentials

The grand mean ERPs to the critical verbs are presented in Fig. 2 for the midline sites and in Fig. 3 for the lateral sites. As can be seen in these figures, the critical verbs elicited a negativity peaking at about 200 ms (N1) and a positivity peaking around 250 ms after the critical verb (P2) which were both maximal at the occipital sites. These components were followed by a broad negative wave in the 250–500 ms epoch peaking at about 350 ms, the N400, which was largest at central and posterior sites. It is well known that the N400 component is elicited by each open class word (e.g., Kutas and Van Petten, 1994). Inspection of the waveforms suggested that mean amplitude was more negative for implausible verbs than for

Table 2
Mean reaction time (RT) and error percentages (Error) with standard deviations (S.D.), for the plausible, implausible and subject-relative, and object-relative sentences

	Subject relative				Object relative				RT		Error	
	RT	S.D.	Error	S.D.	RT	S.D.	Error	S.D.	Mean	S.D.	Mean	S.D.
Plausible	532	29	6.52	.01	587	25	12.08	.02	560	25	9.30	.01
Implausible	502	25	3.97	.01	552	27	8.11	.01	527	24	6.04	.01
Mean	517	26	5.25	.01	570	23	10.1	.01				

The means are marginal means averaged over either complexity or plausibility.

plausible verbs at some midline sites (Fza, Fz, Cz, and Pz) and right hemisphere sites (RAT, RT, RTP).

The N400 was followed by a slow positive shift, the P600, starting at about 600 ms and extending up to 1000 ms which was largest at central and posterior sites. Inspections of the waveforms suggests the presence of a small P600 effect at the midline (see, for example, Pz and Oz) and at some lateral sites of the right hemisphere (see, for example, P4P or OR). The differences between conditions were, however, rather small (about 1 μV or less).

3.3. Statistical analyses

About 11% of the trials were excluded from the analyses because of artifacts; of which 2% belonged to the SR-plausible condition, 2% belonged to the SR-implausible condition, 4% belonged to the OR-plausible condition, and 3% belonged to the OR-implausible condition.

3.3.1. N400 window (400–500 ms)

For the midline sites, the analyses for the N400 window did not reveal a main effect of plausibility, $F < 3$. No main effect of complexity, $F < 1$ or two-way interactions between plausibility and complexity were obtained, Fs < 1. No other interactions with plausibility were observed, all Fs < 2. For the lateral sites, no plausibility effects were obtained, $F < 1$. No main effect of complexity or interactions with plausibility and complexity were obtained, $F < 1$. In addition, no interactions of plausibility with site, ROI and/or hemisphere were obtained, all Fs < 3.5. Additional MANOVAs for the midline and the lateral sites in which N400 was measured in a broader latency window (300–500 ms following verb onset) revealed essentially the same pattern of results, in that no effect of plausibility or relevant interactions were obtained, all Fs < 3.[4]

In sum, the N400 analyses indicated that no N400 effect was observed, neither at the midline nor at the lateral sites. Therefore, no evidence for an N400 plausibility effect was obtained in the present experiment.

3.3.2. P600 window (650–850 ms)

For the midline sites no effect of plausibility, $F < 1$ or effect of complexity was observed, $F < 2$. In addition, no plausibility by site or other interactions with plausibility or complexity were obtained, all Fs < 1. In other words, there were no indications for a P600 effect at the midline sites. For the lateral sites, no main effect of plausibility was obtained, $F < 2$. However, an interaction of plausibility by ROI was found, $F\ (1,35) = 5.89$, $p < .03$. Separate analyses for the two regions of interests indicated that a P600 effect was present at posterior sites, $F\ (1,35) = 6.41$, $p < .02$, but not at anterior sites, $F < 1$. Moreover, a trend for a four-way interaction between plausibility, ROI, hemisphere and site was found, $F\ (3,33) = 2.54$, $p < .08$. To further examine the scalp distribution of the P600 effect post hoc Newman-Keuls' tests were conducted. These tests indicated that a P600 effect was present at the following four sites of the right hemisphere (P4, T6, P4P, OR: $p < .05$) and one single site of the left hemisphere (P3P, $p < .05$).[5]

In sum, the analyses for the P600 for the midline sites demonstrated that at centroparietal sites (Cz and Pz) which are the sites that typically show largest P600 effects to syntactic anomalies, no P600 effect was present. However, the analyses for the lateral sites indicated that the P600 effect was not totally eliminated; since a P600 effect was still obtained at some posterior sites. Thus, although in the present study there was no P600 effect for the midline sites, the latter (lateral) analyses suggest that subjects do not have complete control over their processing strategies.

The results for the N400 are similar to those of the study of Kolk et al. (2003), in that no reliable N400 effect was found. The results for the P600 on the other hand, differ from those of Kolk et al. (2003), because in that study clear P600 effects were obtained, both for the midline and for the lateral sites of the left and right hemisphere. The P600 effect in the present study, therefore, was mainly limited to right posterior sites.

3.3.3. Global analyses

To directly compare the P600 effects, supplementary global analyses for P600, with instruction (semantic plausibility judgment versus focus-on-syntax) as between-subject factor and plausibility (plausible versus implausible) as within-subject factor were carried out. The relevant question was whether a plausibility by instruction interaction would be obtained. The main results of these global analyses were as follows: neither for the midline nor for the lateral sites an effect of instruction and/or an instruction by plausibility interaction was present, Fs < 1.5. Furthermore, no interaction of instruction with site and/or hemisphere was obtained that pointed to a difference in P600 pattern between experiments.

One could argue that given that the RTs in the present study were faster than those in the Kolk et al. study, that the use of the same broad window to quantify P600 effects might not be the best window to capture P600 differences between the two studies.[6]

[4] ERP grand mean inspection suggests that an N400 is present in the 250–450 ms interval for some electrodes of the midline and the right hemisphere. However, statistical analyses in this latency window reveal that these effects are not reliable. We decided to present the results for the 400–500 ms interval in the main text, to be able to directly compare our data set with the data set of Kolk et al. (2003).

[5] ERP grand mean inspection suggests that a P600 is present in the 550–750 ms interval for some midline and posterior electrodes. However, statistical analyses in this latency window reveal essentially the same pattern of results as was revealed in the 650–850 ms window. We decided to present the results for the 650–850 ms interval in the main text, to be able to directly compare our data set with the data set of Kolk et al. (2003).

[6] To examine more closely the onsets and lengths of the differences in the ERP plausibility effect between the two studies, supplementary time course analyses were conducted both for the 300–500 ms window (to capture N400) and the 500–900 ms window (to capture P600) using consecutive bins of 50 ms. To reduce the chance of Type I errors due to the large number of comparisons, an effect is referred to as significant only if it was present in at least two consecutive time epochs. For both kinds of analyses the mean amplitudes were entered into a MANOVA. These analyses showed that neither for the N400 nor for the P600 earlier or more transient reliable differences between the two studies were present.

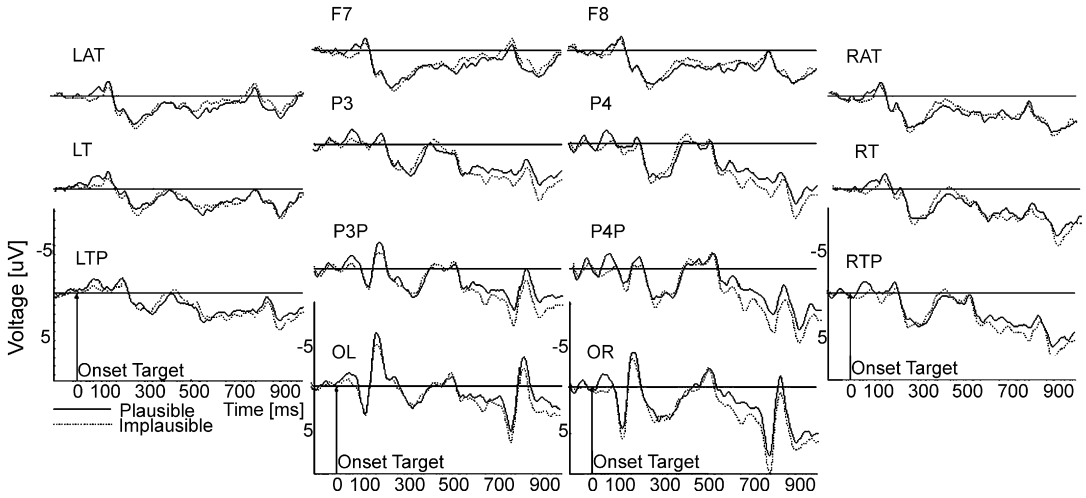

Fig. 3. Grand ERP averages to the critical verb for a representative subset of lateral sites, collapsed over the two levels of complexity. Averages are time locked to the onset of the critical verb, and superimposed for the two levels of Plausibility. Negativity is plotted upwards.

Therefore, supplementary analyses for the P600 peak amplitude, were carried out across studies. To this aim, the P600 peak amplitude was measured between 500 ms and 900 ms after word onset. The factors of the global analyses were: instruction (semantic plausibility judgment versus focus-on-syntax) as between-subject factor and plausibility (plausible versus implausible) as within-subject factor. The main results of the analyses for P600 peak amplitude were as follows: for the lateral sites, no effect of instruction and/or an instruction by plausibility interaction was present, Fs < 1.5. Furthermore, no interaction of instruction with site and/or hemisphere was obtained that pointed at differences in P600 pattern for the lateral sites between experiments.

However, most important for our present purposes, the analyses for the peak amplitude for the midline sites confirmed that there were indeed differences in P600 pattern between the two experiments. For the midline sites, the interaction between plausibility and instruction was significant (F (1,74) = 5.37, $p < .05$). Separate analyses for the two experiments revealed that a P600 effect was present in the study of Kolk et al., F (1,39) = 6.46, $p < .01$, but not in the present study, $F < 1.5$.

Taken together the results from the present experiment and those from the global analyses support the following findings: the focus-on-syntax instruction was successful in reducing the P600 effect. This was reflected by: first, the fact that no P600 effect was present at the midline. The global analyses confirmed a change in P600 pattern as a function of instruction: in particular, the analyses for the peak amplitude showed that a P600 effect was obtained for the midline sites in the study of Kolk et al. (2003) but was absent in the present experiment. Second, opposite to the Kolk et al. (2003) data, the analyses for the lateral sites showed that the P600 effect in the present study was not bilaterally distributed, but mainly limited to a subset of posterior sites over the right hemisphere. Thus, the P600 effect did not totally disappear, as it was still present for a small number of posterior sites. In other words, the change in instruction led to a significant decrease in P600 effect. However, it did not completely abolish the P600 effect.

4. Discussion

It was argued in Section 1 that the language comprehension system uses a combination of heuristics and algorithms (Townsend and Bever, 2001; Ferreira, 2003). Usually, these heuristics and syntactic algorithms produce the same thematic interpretations. Kolk et al. (2003) studied sentences in which these two routes produced a conflicting result; semantic reversal anomalies were used which were formed by exchanging the subject and object of semantically acceptable sentences. They found a P600 effect, and not an N400 effect in response to these semantic reversal anomalies. Kolk et al. (2003) argued that the two routes produce a mismatch between the semantically plausible thematic interpretation proposed by the semantic heuristic which is highly expected based on our world knowledge and the implausible one proposed by the algorithmic parser. This mismatch brings the brain to re-attend the unexpected element to check upon its veridicality. Kolk et al. (2003) proposed that this control operation underlies the P600 effect to semantic reversal anomalies. Since the word-level integration process does not encounter difficulties when integrating the words of a reversal anomaly into a coherent meaning, no N400 effect was obtained.

The present study tested the hypothesis that the P600 effect to semantic reversal anomalies is a reflection of the control operation triggered by the mismatch between thematic interpretations. The focus-on-syntax instruction created a context for participants wherein anomalous sentences were expected; they were explicitly informed that semantic reversals had been constructed on purpose and that they should not be misled by word meanings (‘knowledge of what normally happens in the world’), but attend to the structure of the sentences (‘who does what to whom?’). This instruction should reduce the discrepancy between the thematic interpretations proposed by the semantic heuristic and the syntactic parse. Consequently, semantic reversals were expected and therefore less readily qualified as a possible processing error. So, the necessity of the brain to re-attend the event to check for a

possible processing error should be reduced. If the proposal by Kolk et al. (2003) is right, then our focus-on-syntax instruction should lead to a reduction or disappearance of the P600 effect to semantic reversal anomalies.

The major result of the present article is that our focus-on-syntax instruction did influence both the behavioral data and the ERP data. First, participants in the present study were faster and more accurate than the participants in the Kolk et al.'s (2003) study. This improvement in performance is taken to indicate that the instruction to focus on syntax and not on word meaning was effective, in that our participants were less easily misled by the semantic reversals.

Let us now turn to the ERP effects. The analyses for the midline sites revealed that there were no indications for a P600 effect at the midline sites; which are the sites that typically show the largest P600 effects to syntactic violations. In addition, this result stands in sharp contrast to that of Kolk et al. (2003), who reported a plausibility effect for the midline sites. The global analyses for P600 amplitudes further supported this difference in P600 pattern as a function of instruction. Importantly, an interaction between plausibility and instruction was obtained for the midline sites. Follow up tests verified that this interaction reflected the presence of a P600 effect in the Kolk et al. study ($p < .05$) but absence of this effect in the present study ($F < 1.5$). In addition, at lateral sites the P600 in the present study was reduced in that it was less broadly distributed than the typical syntactic P600 effect. A two-way interaction of plausibility by ROI revealed that the P600 effect was not totally eliminated in the present study, but that a P600 effect was still present for a small set of posterior sites over the right hemisphere. This also stands in contrast to the results of Kolk et al. (2003), who reported a main effect of plausibility for the lateral sites of both the left and right hemisphere.

The focus-on-syntax instruction thus led to a disappearance of the P600 effect at the midline sites and at all but one site of the left hemisphere. We propose that this instruction directed the participants' attention to the structure of the sentences and created a context wherein semantic reversals were expected. The mismatch in the thematic interpretation proposed by the heuristic and the one proposed by the parser was expected and therefore not qualified as a likely processing error. Because the instruction turned an unexpected real life event (that foxes hunt poachers) into a less unexpected event, there was less need for the brain to re-attend the implausible linguistic unit which resulted in a decrement of the P600 effect. Hence, the results of the present study support the proposal by Kolk et al. (2003) that the P600 effect to semantic reversal anomalies is based on a control operation triggered by a mismatch in thematic interpretations.

The fact that a residual P600 effect was still obtained for a small set of posterior sites suggests that the reversal anomalies did to some extent still elicit a mismatch in the thematic interpretation proposed by the semantic heuristic and the one proposed by the syntactic algorithm. Apparently, participants did not have complete control over their natural tendency to give priority to semantic processes. In spite of the focus-on-syntax instruction, the bias for the semantically most plausible interpretation continued to be active, though to a lesser degree. What is critical for the current purposes is that this P600 effect was less broadly distributed than the P600 effect observed in the Kolk et al.'s study (2003), where the same semantic reversals had been presented. This suggests that a substantial part of the P600 effect after semantic reversals can be accounted for by the control operation that is triggered by a mismatch between two thematic interpretations.

The observed reduction in P600 effect is consistent with previous studies that have shown that the P600 effect is affected by task demands and is therefore assumed to reflect a process that is largely under the participant's control (Gunter and Friederici, 1999; Gunter et al., 1997; Hahne and Friederici, 1997). For example, the amplitude of the P600 has been shown to be modulated by probability with larger P600 effects to the less probable event (Coulson et al., 1998; Hahne and Friederici, 1999; Gunter et al., 1997). This indicates that the P600 effect is sensitive to list composition that affects subjects' expectations.

As shown by previous studies, the P600 effect to syntactic and certain semantic violations shows a central/posterior scalp distribution (Coulson et al., 1998; Kolk et al., 2003). The residual P600 in the present experiment (after a focus-on-syntax instruction) had a slightly more posterior scalp distribution, including the right occipital site. Note that also in the Kolk et al. study significant P600 effects were present at occipital sites. At least for the data presented by Kolk et al. (2003), which allowed a direct within-subject comparison between the topography of the P600 after semantic reversal anomalies and the syntactic violations, there were no indications for topographical differences in P600 effect. In the present study, such a within-subject comparison is impossible. Based on the Kolk et al. study, though, we consider the possibility that the present (semantic) P600 effect is qualitatively different in terms of scalp distribution as rather unlikely. Furthermore, we would like to point out that there is evidence for some variation in scalp distribution of the syntactic P600 effect. A more frontal/broad distribution of the P600 effect has, for example, been reported for locally ambiguous sentences (Friederici et al., 1996; Hagoort et al., 1999; Osterhout and Holcomb, 1992; Van Berkum et al., 1999) while Kaan and Swaab (2003) observed a more posterior distribution.

Bever et al. (1998) propose that the semantic interpretation developed by the heuristic depends on passively accumulated and applied statistical generalizations. Thus, according to this proposal, the initial semantic analysis proceeds on the basis of statistically sensitive perceptual strategies. As described in Section 1, the most strongly confirmed abstract pattern available to the heuristic is the word-order strategy. This implies that a noun phrase preceding and agreeing in number with a verb is taken as the subject of that verb. Similarly, a noun phrase following a verb is taken to be the object of the verb. However, this proposal was not supported in the present study. If participants would assign this 'favoured' canonical form to our semantic reversal sentences, they would not have been misled by our reversals. In a sentence like ''The fox that hunted[singular] the poachers stalked through the woods'', the noun phrase preceding the verb is the subject of the verb, and the noun phrase

following the verb is the object of the verb. But, the present experiment indicated that participants were, though to a lesser degree compared to the Kolk et al. (2003), misled by the reversals. This is bolstered by the error percentages (participants had higher error percentages for the implausible sentences compared to the plausible sentences) and the residual P600 effects for the implausible compared to the plausible sentences.

No evidence was obtained for an N400 effect in the present study. Consistent with our previous reasoning, this is because the content words of the semantic reversals in the implausible and plausible sentences were comprised of the same lexical items that were easy to integrate into a coherent meaning. No integration difficulty and hence no N400 effect was expected to occur. This supports the claim that the semantic heuristic develops a thematic interpretation on the basis of the *meaning* of the individual words. So, participants are not biased to follow the canonical form, but they are biased by the meanings of the individual words and their world knowledge. This is in agreement with Bever's (1970) description of 'Strategy C' and Ferreira's (2003) description of the 'plausibility heuristic'. These heuristics propose a semantic interpretation of an utterance which is in agreement with knowledge of the individual content words or schema's in long term-memory. The use of heuristics has also been shown to play an important role in the explanation of comprehension difficulty by aphasic patients (Saffran et al., 1989).

As presented in Section 1, 'syntax-first' models propose that a modular syntactic processing system guides and precedes semantic interpretation. Semantic information is used only in a second stage, either to choose between different structural possibilities or to guide revision after the chosen structure turned out to be erroneous. The syntactic structure of our semantic reversals is unambiguous. That is, in the reversal sentence *the fox that hunted the poacher*, there is only one option: *fox* is the Agent, and *poacher* is the Theme. A syntax-first model would thus predict a semantically incorrect representation of this sentence. This would predict the modulation of the N400 which, however, was not found. Constraint-based models propose that semantic information is used during syntactic structure build-up. Hence, semantic information may help to choose between different structural possibilities. But, semantic information cannot propose structural possibilities. If syntactic cues are unambiguous, then semantic information does not exert a controlling influence. This implies that the constraint-based models would also predict an N400 effect after our semantic reversals. Hence, both constraint-based models as syntax-first models assume that syntactic information, when unambiguous, will control the initial combinatory analysis of linguistic input. Contrary to what both models predict, we report that semantic information can independently propose a thematic interpretation of a sentence, overwhelming unambiguous syntactic cues. The data from our laboratory seem to be more consistent with a language processing system of parallel, independent syntactic and semantic processing mechanisms; in which semantic processing can overrule unambiguous syntactic cues (see also, Kim and Osterhout, 2005).

In conclusion, our results support the idea that the P600 effect to semantic reversals is based on a control operation after a mismatch in thematic interpretations to check upon the possibility of a processing error. The present results show that the focus-on-syntax instruction recruited the neural system underlying the control operation to a lesser degree than the semantic plausibility judgment instruction in the experiment by Kolk et al. (2003). This is demonstrated by the disappearance of the P600 effect most notably at central and parietal midline sites (which typically show maximal P600 effects) and the fact that the P600 effect was not bilaterally distributed but mainly limited to the posterior areas over the right hemisphere. However, the residual P600 effects indicate that the focus-on-syntax instruction did not completely resolve the conflict between the thematic interpretations proposed by the semantic heuristic and the syntactic parse as a possible reading error. This suggests that the semantic heuristic and hence the control operation to check for a processing error continued to be active to some extent. In the face of strong expectations which are incongruent with the syntactically driven interpretations, it may be difficult to block the monitoring response completely.

Acknowledgments

Portions of this research were presented at the 11th Annual Meeting of Cognitive Neuroscience Society in San Francisco, 2004. We thank two anonymous reviewers for their very helpful comments on an earlier version of this article. We are grateful to the ERG group of the NICI for technical assistance and to Maarten vd Meulen for his support.

References

Besson, M., Fischler, I., Boaz, T., Raney, G., 1992. Effects of automatic associative activation on explicit and implicit memory tests. Journal of Experimental Psychology: Learning, Memory, and Cognition 18, 89–105.

Bever, T.G., 1970. The cognitive basis for linguistic structures. In: Hayes, J.R. (Ed.), Cognition and the Development of Language. Wiley, New York, pp. 279–362.

Bever, T.G., Sanz, M., Townsend, D.J., 1998. The emperor's psycholinguistics. Journal of Psycholinguistic Research 27 (2), 261–284.

Chomsky, N., 1957. Syntactic Structures. Mouton & Co., The Hague/Paris.

Chwilla, D.J., Brown, C.M., Hagoort, P., 1995. The N400 as a function of the level of processing. Psychophysiology 32 (3), 274–285.

Chwilla, D.J., Hagoort, P., Brown, C.M., 1998. The mechanism underlying backward priming in a lexical decision task: spreading activation versus semantic matching. Quarterly Journal of Experimental Psychology: Human Experimental Psychology 51A (3), 531–560.

Chwilla, D.J., Kolk, H.H.J., Mulder, G., 2000. Mediated priming in the lexical decision task: evidence from event-related potentials and reaction time. Journal of Memory and Language 42 (3), 314–341.

Coulson, S., King, J.W., Kutas, M., 1998. Expect the unexpected: event-related brain response to morphosyntactic violations. Language and Cognitive Processes 13 (1), 21–58.

Craik, F.I.M., Lockhart, R.S., 1972. Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behavior 11, 671–684.

Dien, J., Santuzzi, A.M., 2004. Application of repeated measures ANOVA to high-density ERP datasets: a review and tutorial. In: Handy, T.C. (Ed.), Event-Related Potentials: A Methods Book. The MIT Press, Cambridge, Massachusetts, pp. 57–82.

Ferreira, F., 2003. The misinterpretation of noncanonical sentences. Cognitive Psychology 47 (2), 164–203.

Ferreira, F., Bailey, K.G.D., Ferraro, V., 2002. Good-enough representations in language comprehension. Current Directions in Psychological Science 11 (1), 11–15.

Ferreira, F., Clifton, C., 1986. The independence of syntactic processing. Journal of Memory and Language 25 (3), 348–368.

Fodor, J.A., 1983. The Modularity of Mind. Bradford Books, Cambridge, MA.

Frazier, L., Fodor, J.D., 1978. The sausage machine: a new two-stage parsing model. Cognition 6 (4), 291–325.

Friederici, A.D., Hahne, A., Mecklinger, A., 1996. Temporal structure of syntactic parsing: early and late event-related brain potential effects. Journal of Experimental Psychology: Learning, Memory, and Cognition 22, 1219–1248.

Gigerenzer, G., Goldstein, D.G., 1996. Reasoning the fast and frugal way: models of bounded rationality. Psychological Review 103 (4), 650–669.

Gigerenzer, G., ABC Research Group, 1999. Simple Heuristics that Make Us Smart. Oxford University Press, New York.

Gunter, T.C., Friederici, A.D., 1999. Concerning the automaticity of syntactic processing. Psychophysiology 36 (1), 126–137.

Gunter, T.C., Stowe, L.A., Mulder, G., 1997. When syntax meets semantics. Psychophysiology 34 (6), 660–676.

Hahne, A., Friederici, A.D., 1997. Two stages in parsing: early automatic and late controlled processes. Experimental Brain Research 117, 47.

Hahne, A., Friederici, A.D., 1999. Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. Journal of Cognitive Neuroscience 11 (2), 194–205.

Hagoort, P., Brown, C., Osterhout, L., 1999. The neurocognition of syntactic processing. In: Brown, C.M., Hagoort, P. (Eds.), The Neurocognition of Language. University Press, Oxford, pp. 273–317.

Hoeks, J., Stowe, L.A., Doedens, G., 2004. Seeing words in context: the interaction of lexical and sentence level information during reading, Cognitive. Brain Research 19, 59–73.

Holcomb, P.J., Neville, H.J., 1990. Semantic priming in visual and auditory lexical decision: A between modality comparison. Language and Cognitive Processes 5, 281–312.

Kaan, E., Swaab, T.Y., 2003. Repair, revision, and complexity in syntactic analysis: an electrophysiological differentiation. Journal of Cognitive Neuroscience 15, 98–110.

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

Kolk, H.J., Chwilla, D.J., Van Herten, M., Oor, P.J.W., 2003. Structure and limited capacity in verbal working memory: a study with event related potentials. Brain and language 85 (1), 1–36.

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

Kuperberg, G.R., Caplan, D., Sitnikova, T., Eddy, M., Holcomb, P.J., 2006. Neural correlates of processing syntactic, semantic and thematic relationships in sentences. Language and Cognitive Processes 21, 489–530.

Kutas, M., 1993. In the company of other words: electrophysiological evidence for single-word and sentence context effects. Language and Cognitive Processes 8, 533–572.

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

Kutas, M., Hillyard, S.A., 1989. An electrophysiological probe of incidental semantic association. Journal of Cognitive Neuroscience 1, 38–49.

Kutas, M., Schmitt, B.M., 2003. Language in microvolts. In: Banich, M.T., Mack, M. (Eds.), Mind, Brain, and Language: Multidisciplinary Perspectives. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 171–209.

Kutas, M., Van Petten, C.K.F., 1994. Psycholinguistics electrified. Event-related brain potential investigations. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics. Academic Press, San Diego, CA, pp. 83–143.

McDonald, M.C., Maryellen, C., Pearlmutter, N.J., Seidenberg, M.S., 1994. Lexical nature of syntactic ambiguity resolution. Psychological Review 101 (4), 676–703.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh Inventory. Neuropsychologia 9 (1), 97–113.

Osterhout, L., Holcomb, P.J., 1992. Event-related brain potentials elicited by syntactic anomaly. Journal of Memory and Language 31, 785–806.

Saffran, E.M., Schwartz, M.F., Linebarger, M.C., 1989. Semantic influences on thematic role assignment: evidence from normals and aphasics. Brain and Language 62, 255–297.

St. John, M.F., McClelland, J.L., 1990. Learning and applying contextual constraints in sentence comprehension. Artificial Intelligence 46, 217–257.

Townsend, D.J., Bever, T.G., 2001. Sentence Comprehension: The Integration of Habits and Rules. The MIT Press, Cambridge, MA.

Trueswell, J.C., Tanenhaus, M.K., 1994. Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In: Clifton, C., Rayner, K., Frazier, L. (Eds.), Perspectives on Sentence Processing. Lawrence Erlbaum Associates, Hillsdale, NJ.

Van Berkum, J.J.A., Brown, C.M., Hagoort, P., 1999. Early referential context effects in sentence processing: evidence from event-related brain potentials. Journal of Memory and Language 41, 147–182.

Van Herten, M., Kolk, H.H.J., Chwilla, D.J., 2005. An ERP study of P600 effects elicited by semantic anomalies. Cognitive Brain Research 22, 241–255.

Vasey, M.W., Tayer, J.F., 1987. The continuing problem of false positives in repeated measures ANOVA in psychophysiology: a multivariate solution. Psychophysiology 24, 479–486.

Vissers, C.Th.W.M., Chwilla, D.J., Kolk, H.J., 2006. Monitoring in language perception: the effect of misspellings of words in highly constrained sentences. Brain Research 1106, 150–163.