PROCESSING IT-CLEFT SENTENCES IN BRAZILIAN PORTUGUESE: AN ERP STUDY OF LEFTWARD-MOVED CONSTITUENTS IN ROLE-REVERSED SENTENCES

O PROCESSAMENTO DE SENTENÇAS CLIVADAS NO PORTUGUÊS DO BRASIL: UM ESTUDO DE ERP DE CONSTITUÍNTES MOVIDOS PARA A ESquerda EM SENTENÇAS COM PAPEIS TEMÁTICOS REVERSOS

Juliana Novo Gomes1
Aniela Improta França2

ABSTRACT

In this paper we deepened on the processing of cleft role-reversed structures, based on empirical evidence of standard Brazilian Portuguese (BP). We used the electrophysiological technique (EEG/ERP) to map distinguished syntactic and semantic processes, for instance, the N400 and the P600, addressing the focus structures in It-clefts clauses structured in three different experimental conditions: (i) Congruous Cleft Condition: It was the SURFER that the shark attacked in Hawaii; (ii) Reversed Cleft Condition: It was the SHARK that the surfer attacked in Hawaii; and (iii) Incongruous Cleft Condition: It was the COUCH that the shark attacked in Hawaii. Taken together, our findings suggest that the presence of P600s related to role-reversed sentences in previous studies could be attributed to the syntactic reanalysis, instead of the processing of the role reversed item per se. Also, the presence of an N400 effect to the reversals could be due to the frustration of the strong combination of contextual constraints and strong lexical association. Our results make a unique contribution to the ERP response profiles, specially regarding the relationship between the role-reversals and the animacy violations in the Cleft structural frame. Our ERP findings seem to be compatible with the long-held assumption that the N400 and P600 appear to be modulated by the subject-object asymmetry, and were sensitive to, respectively, the semantic attraction between words in the sentences and, the congruency of the predicate. We thus claim that the syntactic anomalies blocked the detection of semantic anomalies, therefore, semantically incongruous sentences, such as role-reversals were perceived to be odd due to a syntactic constraint satisfaction that assigns the right theta-roles to the verbs arguments despite the semantic cues.

KEYWORDS: ERP; sentence processing, role reversal structures.

RESUMO

Neste artigo nos profundamos no processamento de estruturas clivadas com argumentos reversos, com base nas evidências empíricas do Português Brasileiro (PB) padrão. Utilizamos a técnica eletrofisiológica (EEG / ERP) para mapear processos sintáticos e semânticos distintos, por exemplo, o N400 e o P600, relacionados às estruturas de foco em cláusulas estruturadas em três diferentes condições experimentais: (i) Condição

2 Universidade Federal do Rio de Janeiro (UFRJ/CNPq/FAPERJ). Professora do Programa de Pós-Graduação em Linguística e do Departamento de Linguística e Filologia. Contato: aniela@letras.ufrj.br.
clivada Congruente: Foi o SURFISTA que o tubarão atacou no Havaí; (ii) Condição de clivada Invertida: Foi o TUBARÃO que o surfista atacou no Havaí; e (iii) Condição clivada incongruente: Foi o SOFÁ que o tubarão atacou no Havaí. Tomados em conjunto, nossos resultados sugerem que a presença de P600s relacionados a sentenças invertidas em estudos anteriores poderia ser atribuída à reanálise sintática, em vez do ser relativo ao processamento do item de papel invertido em si. Além disso, a presença de um efeito N400 nas reversões de argumento pode ser devido à frustração da forte combinação de restrições contextuais combinadas a forte associação lexical. Nossos resultados prestam alguma contribuição para o entendimento dos padrões de resposta do ERP, especialmente no que diz respeito à relação entre as reversões de papéis e as violações de animacidade no quadro estrutural das clivadas. Os ERP que colhemos parecem ser compatíveis com a suposição já de longa data de que o N400 e P600 parecem ser modulados pela assimetria sujeito-objeto, e são também sensíveis à atração semântica entre as palavras nas frases. Assim, afirmamos que as anomalias sintáticas bloquearam a detecção de anomalias semânticas, portanto, sentenças semanticamente incongruentes, como reversões de papéis, foram percebidas como estranhas devido a uma satisfação de restrição sintática que atribui os papéis teta corretos aos argumentos dos verbos, apesar do viés semântica.

**PALAVRAS CHAVE**: ERP, processamento de sentença, estruturas reversas.

### 1. INTRODUCTION

Several theoretical studies on focus constructions have attracted a considerable amount of attention, because of their syntactic diversity among natural languages. Focus can be realized in different language-specific ways, and it is common to have more than one way of encoding focus in a language (LAMBRECHT, 1994; 2001; LAMBRECHT & POLINSKY, 1997; KISS, 1998; REEVE, 2011, BRAGA et al 2013). Among them, a focus construction that is particularly interesting is the so-called “It -cleft” (ROSS, 1986; 1997; 2000, BRAGA et al 2013). For example, consider the English and Portuguese It-cleft constructions shown in (1a-b).

(1)

a. Subject cleft constructions (SC):

ENG: It was the BEAR that <gap> devoured the poacher in the forest.

PORT: Foi o URSO que <gap> devorou o caçador na floresta.

b. Object cleft constructions (OC):

ENG: It was the POACHER that <gap> the bear devoured in the forest.

PORT: Foi o CAÇADOR que <gap> o urso devorou na floresta.
In such constructions focus may be marked via particular prosodic contours, as in English, and in Brazilian Portuguese, where the pitch accent (indicated by capital letters) falls on bear/urso and poacher/caçador (for encompassing reviews on Brazilian focus structures, see KATO, RAPOSO, 1996, BARBOSA, 2000, 2013; KATO, RIBEIRO, 2009; KATO, 2015; BRAGA et al 2013). Focus is mostly defined as the part of an utterance that introduces new or newly activated information into the current discourse and can be defined operationally as a well-formed answer to a Wh-question (PRINCE, 1978; LAMBRECHT, 1994; ROCHEMONT, 1986; SELKIRK, 1995; KATO, MIOTO, 2005; HIGGINS, 2015). Syntactically, a focused construction encompasses a subject that is modified by a free relative clause or simply a relative clause in the subject position (That [which] attacked the surfer at the beach/ que [que] atacou o surfista na praia) and a predicate phrase (was the shark/ foi o tubarão). The It-cleft version of this sentence is similarly divided, with the same constituency; however in this case the predicate phrase (It is X/ Ser X) precedes the relative clause subject (It was the bear [that devoured the poacher] / Foi o urso [que devorou o caçador]) in the linear order of the sentence (CASTELEIROS, 1979; ROSS, 1986; COSTA AND DUARTE, 2001; LOBO, 2006; MODESTO, 1995, 2001; COWLES ET AL., 2007; KATO, RIBEIRO, 2009; BARBOSA, 2013; LOBO; SANTOS; SOARES-JESEL, 2016.).

Exploring focus constructions empirically is specially significant because often a string of words can be given more than one syntactic analysis, specially when they are formed through movement. Moreover, there are asymmetries in the production and processing of subject clefts vs. object clefts – subject clefts, such as in (1a), are more frequently produced and easily processed than object clefts, such as in (1b), both in spontaneous production and in elicited tasks (COSTA, LOBO, SILVA, 2011; KAHRAMAN, 2011; LOBO, SANTOS, SOARES-JESEL, 2016; DE VINCENZI, 1991; 1996; GOUVÊA, 2005; FANSELOW et al, 2005; HOFMEISTER ET. AL., 2007; KWON et al., 2010). These asymmetries have been attributed to a specific processing difficulty with structures that involve intervention, as proposed by Friedmann, Belletti and Rizzi (2009).

Taken either as a result of the application of Parsing Principles, such as, the Minimal Attachment Principle (CLIFTON & FRAZIER, 1989) or the Minimal Chain Principle (DE VINCENZI, 1991, 1996), the studies under the subject-object asymmetry measuring greater processing complexity for the object integration have enriched the debate on the syntactic processing of several constructions, as well as the reflection on the nature of the processes that
determine the linguistic development and processing.

Nevertheless, resorting to linguistic theory, it is well established that thematic roles are hierarchically ordered with respect to one another (BRESNAN & KANERVA, 1989; Bresnan, 2000; GIVON, 1984; JACKENDOFF, 1972; PRIMUS, 1998; 1999; GOMES, 2009) and that, in some situations, the correct assignment of thematic role to the subject can only be done in a principled way after the full computation of the object.

In fact, many of such predictions have been tested in neuroscience experiments. There are electrophysiological findings, using EEG/ERP methodology, testing how processing operates in terms of hierarchizing processes, drawing upon underlying morphological case information and not on surface order (BADER & LASSER, 1994; KAMIDE & MITCHELL, 1999; FRISCH, 2000; FRISCH & SCHLESEWSKY, 2001; BORNKESSEL ET AL., 2003; GOMES, 2009; 2013; 2014; GOMES & FRANÇA, 2008; 2013; 2015; GOMES et al, 2017).

Such experiments were specially based on the detection of a special wave – N400 – a negative going, plotted downward per convention, that can be sensed at the scalp of the participant 400ms after the presentation of the stimulus. The technique that separates this specific wave out from many other cortical waves that result from other cognitive processes happening together with the linguistic stimulation is named. An incongruous word in its syntactic context, for instance an incongruous object for the eventful requirements of the verb (eat *sandals*), elicits a large amplitude N400, while a congruous one (eat *sandwich*) elicits a small amplitude N400. This is a pretty robust effect shown in over 100 languages.

After 2000, these ERP tests have been amply applied to the subject-object asymmetry field, specially testing role-reversed sentences (HOEKS, STOWE, DOEDENS, 2004; KIM; OSTERHOUT, 2005; KOLK ET AL, 2003; KUPERBERG ET AL, 2003A, 2006, 2007; LAU ET AL, 2006; STROUD; COLIN PHILLIPS, 2010; VAN HERTEN; KOLK; CHWILLA, 2005A; VISSERS; CHWILLA; KOLK, 2007; GOMES, 2014; GOMES & FRANÇA, 2015). For a gist of the materials tested in these experiments take for instance, 2a and 2b.

2a. The bear devoured the poacher in the forest.
2b. The poacher devoured the bear in the forest

---

3 In “Paul broke his vase” vs “Paul broke his arm”, the subject can only be interpreted respectively as agent and experiencer after the full computation of the object.
4 ERP (Event-related Brain Potential)
Gomes (2014) showed that in 2a, a regular order sentence, when electrophysiological activity is measured at the presentation of the object, *the poacher*, a small amplitude N400 appears, signaling the low cognitive demand for the merge that integrates verb and object. *The poacher* is an unlikely complement for *devour*. In 2b, after a larger amplitude N400, since *a bear* is an even more unlikely complement for *devour*, there is a P600 signaling the need of further computation. This P600 is not found in 2a. Notice that the finding of a larger amplitude N400 in 2b coincides with literature testifying that the N400 is a precise congruence marker that is sensitive to different degrees of incongruity.

Let’s examine 3a, 3b and 3c:

3a. She was stung by a fly
3b. She was stung by a wasp
3c. She was stung by a bee.

Kutas & Hillyard (1984) point to the fact that the presentation of the prepositional phrase in 3a elicited the largest amplitude N400 of the three, since *sting* does not select *fly* for complement. The second largest amplitude, medium amplitude N400, is yielded by 3b, and a small amplitude one by 3c.

Thus, at the point of the complement a large amplitude N400 is detected due to the unexpected complements (*2a devoured the poacher vs 2b devoured the bear*) both non default edible objects, but larger for 2b. So up to that point the two sentences elicit similar ERPs.

Finally, the incongruity in 2b kicks in at 600 ms. So, it happens after the N400 eliciting the vP closure. In 2b there is indication that the subject computation underwent a late and costly process that lasted until after the object was considered and found not fitting. This piecemeal analysis, of the different processing points, provided by the ERP technology, highlighted asymmetries between the subject and the object integration, making us argue that the integration of the subject is more costly than that of the object and depends on it.

Thus, the discrepancy in relation to the findings of psycholinguistic experiments, vouching for a syntactic asymmetry that places more difficulty on the object integration, and that of electrophysiological studies, that detect late effects of subject computation well after object computation, calls for further investigation. Cleft sentences for their structure can offer other points of observation to these facts. In fact, in this article, we will investigate the processing of It-Cleft
sentences, comparing Object It-Cleft sentences, in three different conditions: congruous, reversed and incongruous, as in (4), (5) and (6):

(4) “It was the SURFER that the shark attacked in Hawaii”

(PORT) Foi o SURFISTA que o tubarão atacou no Havaí.

(5) “It was the SHARK that the surfer attacked in Hawaii”

(PORT) Foi o TUBARÃO que o surfista atacou no Havaí.

(6) “It was the COUCH that the shark attacked in Hawaii”

(PORT) Foi o SOFA que o tubarão atacou no Havaí.

2. CLEFT CONSTRUCTIONS

One of the first studies on cleft sentences appeared in Chomsky (1977), where cleft sentences were analyzed in the same fashion as relative clauses: the focused XP functioned as the head of the relative clause or relative-like clause with an A-bar gap\(^5\) (CHOMSKY, 1977). This gap is co-indexed with the predicative argument. The focus XP was base-generated and followed by a phonetically null operator. This operator was moved from an argument position inside the relative clause, leaving a trace behind at the extraction site inside the relative clause. Thus, the null operator analysis of cleft structures is much the same as that of Wh- movement in relative clauses as seen in (4) and (5) (For more on the cleft sentences analysis see: BARBOSA, 2013).

\(^5\) Clefs pass standard diagnostics of A-bar movement, including island sensitivity, weak crossover effects, parasitic gap licensing, etc. (e.g. Chomsky 1977).

\textit{It}-Cleft clauses, such as (4) to (5), as Wh-questions, appear in sentence-initial position, displaced from their canonical argument position. It is well known that when speakers process those kinds of sentences, they try to associate the sentence-initial phrase with the verb in the main clause, meaning that they attempt to locate the gap position, from which the filler-phrase was displaced. The relationship between the surface position of the filler-phrase and the gap is known as a filler-gap dependency. Thus, according to Phillips (2003), when speakers incrementally process a sentence like (4), they initially interpret the filler-phrase as the object of the main verb, and are then forced to revise this analysis when they encounter the complementizer ‘that’, which signals the presence of an embedded clause. The initial-phrase is ultimately interpreted as the subject of the embedded verb ‘attacked.'
The generalization that speakers attempt to posit a gap for the fronted phrase at the first available opportunity is well established, and has been documented in a number of languages (FRAZIER & FLORES D’ARCAIS, 1989; DE VINCENZI, 1991; SCHLESEWSKY et al., 2000). In order to capture this generalization, Frazier and colleagues proposed that there is a parsing-specific strategy that leads the parser to create a gap for a displaced phrase in the first possible position (‘Active Filler Strategy’: Frazier, 1987b; de Vincenzi, 1991). An alternative approach to this generalization argues that there is no specialized strategy for creating filler-gap dependencies. Rather, the observed effects arise as a direct consequence of the need to satisfy thematic role assignment constraints as soon as possible (Pritchett, 1992; Gibson, Hickok, & Schütze, 1994). Thus, an It-Cleft clause is preferably associated with a gap position in the main clause, because that is the first position where thematic role assignment is possible. Thus, see (7) below.

(7)

a. Subject cleft construction (SC):

It was the shark\textsubscript{1} that t\textsubscript{1} attacked the surfer at the beach.

Foi o tubarão\textsubscript{1} que t\textsubscript{1} atacou o surfista na praia.

b. Object cleft constructions (OC):

It was the surfer\textsubscript{1} that the shark attacked t\textsubscript{1} at the beach.

Foi o surfista\textsubscript{1} que o tubarão t\textsubscript{1} atacou na praia.

In (7a) we can see a Subject It-Cleft clause both in English and Portuguese, in which the displaced element (the shark\textsubscript{1}/o tubarão\textsubscript{1}) is the subject of the relative clause (t\textsubscript{1} attacked the surfer/t\textsubscript{1} atacou o surfista). As for in (7b), we illustrate an Object It-Cleft clause, in which the object (the surfer\textsubscript{1}/o surfista\textsubscript{1}) is the displaced element of the relative clause (the shark attacked t\textsubscript{1}/o tubarão atacou t\textsubscript{1}).

Earlier experimental studies investigating It-Cleft Constructions have shown that subject-extracted constructions are easier to process than object-extracted constructions in a variety of languages (see i.e., in English: FERREIRA, 2003; GORDON, et al 2001; in Japanese: KAHRAMAN et al, 2011a and b; MASATAKA, TATEYAMA, SAKAMOTO, 2015; in Korean: KWON, NAYOUNG, 2008; in French: ROBERT, 2014).
Two hypotheses have been proposed to account for this subject-gap preference in English, namely the “Structural Distance Hypothesis (SDH)” (O’Grady 1997) and the “Dependency Locality Theory (DLT)”, often called the “Linear Distance Hypothesis (LDH)” (Gibson 1998, 2000). Both hypotheses assume that one of the major sources of processing difficulties in cleft constructions is the integration of the displaced element (i.e., a filler, such as the shark/surfer in (7a-b)) into its original position (i.e., a gap). However, as their names suggest, they use different distance metrics.

The SDH hypothesis (O’Grady 1997) predicts that, in all languages, subject-extracted constructions will be easier to process than object-extracted constructions because subjects are hierarchically higher and are closer to the gap position than objects.

This account has been experimentally tested (Gibson and Thomas, 1999), but some authors have argued that several factors other than metrical ones may affect the relative difficulties in gap-filling parsing (Cowles et al., 2007).

One possible factor has to do with the expectancy of a canonical word order. A study investigating the influence of a preceding Wh-question on canonical preferred (subject-object-verb [SOV]) versus non-preferred (object-subject-verb [OSV]) word order options in German (Bornkessel, Schlesewsky & Friederici, 2003) manipulated a wide variety of Wh-question contexts (in which case marking and word order were manipulated relative to both SOV and OSV target sentences). The results suggest that expectations such as these about general information delivery parameters of word order influence on the faster recognition of a structure. Thus, it is safe to assume that speakers will find the most frequent word order structure in a specific language the easiest and, consequently, it will be the fastest one to be processed.

Ferreira (2003) has argued that especially when a sentence becomes extremely difficult to parse and interpret, comprehenders could adopt a good-enough strategy to trying to understand it. Sentences like (8c) and (8d) when compared with (8a) seem to fit the “extremely difficult to parse and interpret” scenario proposed by Ferreira.

(8) What attacked the surfer at the beach?
   a. * It was the shark1 that t1 attacked the surfer at the beach.
   b. # It was the surfer1 that the shark attacked t1 at the beach.
   c. * It was the shark1 that the surfer attacked t1 at the beach.
   d. * It was the surfer1 that t1 attacked the shark at the beach.
As argued previously with examples in (8), when answering about an agent of a surfer-attack, the speaker is expecting an agent in the focus of the answer, as in (8a). Likewise, as in (8b), (8b) is not incongruous, but is infelicitous (marked with #). Furthermore, when a speaker encounters (8c), the information focused in the It-cleft position is a good fit as the agent of a surfer-attacking, but as the speaker unveils the sentence, it turns out that the It-cleft position is actually filled with the object of the predicate of the surfer attack, meaning that the surfer attacked the shark. Thus, the comprehension system would be expecting the shark as the doer of the attack and the surfer as the patient of it, semantically speaking. Moreover, syntactically, the It-Cleft sets up the expectation that the cleft position will be dedicated to the agent of the event (attack), as requested in the preceding Wh- question in (8), but instead, in (8c) this position contains a noun phrase that is animate, but implausible as the patient of a surfer-attack.

Similarly but reversely, in (8d), the comprehension system would be expecting surfer as the patient of a shark-attack and the shark as the agent of it semantically speaking, and syntactically, the agent of the attack would be at the cleft position. But as the sentence unveiled itself, in (8d) this position contains a noun phrase that is also animate, but implausible as the agent of a shark-attack.

It might be extremely difficult to either overpass the focus information in the It-cleft position, or to recover from the first commitment, in (8c) and in (8d), for different reasons. In (8c) ‘the shark’ is a highly plausible agent for the surfer-attack, and also a suitable answer for the question in (8). Thus, (8c) is a perfect scenario to put in test the ‘good enough’ approach, since it could be the case that the speaker would take the clefted information (the shark) as the agent/subject of the sentence and adopting a ‘good enough’ approach, attribute the theme role to the predicate (the surfer), contradicting the syntactic structure without further analysis.

On the other hand, the sentence in (8d) starts with an odd focused information (the surfer) as an agent of an attack. Remember that we are dealing with a question-answer setup. Thus, if the processor is being guided by the semantic cues (theme assignment), ‘the surfer’ is expected to be the ‘patient’ of an attack, not the agent of it, as the syntactic structure implements. Respectively, ‘the shark’ would be a highly plausible agent for the surfer-attack, but not in the syntactic context exemplified in (8d).

The results for these scenarios, suggested by Ferreira and colleagues (2002), are that the speakers would even tolerate outright ungrammaticality, as in (8c) and (8d) (for more on Good-enough representations see FERREIRA, BAILEY, FERRARO, 2002).
Cowles and colleagues (2007) explain this type of mismatch in terms of violations of information structure:

When a noun phrase that has been assigned the thematic role of undergoer in the previous discourse context appears in a syntactic position to which the same discourse context has led the comprehension system to expect that it can assign the thematic role of agent, multiple simultaneous sources of difficulty in thematic role assignment result: the prior assignment of the undergoer role is called into question, the agent role cannot be discharged fully or immediately at the expected syntactic position, and must therefore continue to be held in working memory pending resolution, and the agent of the event remains unidentified both in the current sentence as well as in the overall discourse representation (Cowles, 2007).

As shown in previous studies (FERREIRA, BAILEY, FERRARO, 2002; KOLK, ET AL., 2003; KUPERBERG ET AL., 2003; HOEKS, STOWE, DOEDENS, 2004; KIM, OSTERHOUT, 2005; VAN HERTEN, KOLK, CHWILLA, 2005; WANG, HAGOORT, YANG, 2009; KOS, VOSSE, VAN DEN BRINK, HAGOORT, 2010; WANG ET AL., 2012; GOMES, 2014; GOMES, FRANÇA, 2015) in some specific scenarios characterized as role reversals, the agent/patient thematic positions are reversed, but the sentences can still be perceived as plausible. See examples in (9):

(9) Role-reversed sentences in Active and Passive Voices, as reversed from the canonical ‘the shark attacked the surfer at the beach’ (GOMES, 2014, GOMES e FRANÇA, 2015).

a. The surfer attacked the shark at the beach.

b. The shark was attacked by the surfer at the beach.

Gomes (2014) and Gomes e França (2015) based their experiments on the assumption that the expectations produced by semantic constraints are not necessarily the same as those generated by syntax. On one side, semantic context incrementally restricts the range of possible expectations for upcoming words in a sentence (KUTAS & HILLYARD, 1984; VAN PETTEN & KUTAS, 1990) or discourse (KUTAS & FEDERMEIER, 2000; VAN BERKUM, ZWITSERLOOD, HAGOORT & BROWN, 2003). On the other, syntax generates cognitive processes associated with hierarchical positions.

This is especially relevant for the It-cleft constructions used in this new study, because, the right edge of the predicate constituent (It is X) is the first position, where relevant information is given. As discussed in the previous section, clefts are associated with both idiosyncratic
interpretive properties and non-canonical syntax, but the precise mechanisms for deriving them are still debated\(^6\). Despite that, the *It*-Cleft syntactic device *per se* seems to be asymmetric and difficult to interpret, as it involves displacement and, frequently, it contradicts the word order of its counterpart canonical sentence, as exemplified in (8) and (9) above. Therefore, the *It*-Clefts used in this new study could be considered under the role-reversed sentences as a subject-object asymmetry case.

This subject-object asymmetry encounters some evidence from developmental work (BEVER 1970; LEMPERT AND KINSBOURNE 1980; COSTA, LOBO AND SILVA, 2011; M LOBO, SANTOS, SOARES-JESEL, VAZ, 2014). They report that children have asymmetric difficulties interpreting object clefts compared to subject clefts both in English and in Portuguese. The assumption behind these results is that while both subject and object clefts are non-canonical in employing the more complex cleft syntax, lexical elements in subject *It*-clefts conform to the canonical SVO word-order and were therefore amenable to surface word-order based interpretive strategies. A word-order based strategy would not work in the case of object clefts, where the arguments do not linearly correspond to the canonical word order.

Another approach would come from experimental syntax. Recently results show that the processor may be sensitive not only to parsing difficulties, but also to reversals of pragmatic plausibility relations, as a violation either of thematic relations, or to selectional restrictions on the verb (KUPERBERG, SITNIKOVA, CAPLAN, & HOLCOMB, 2003; KOLK, CHWILLA, VAN HERTEN, & OOR, 2003; HOEKS, STOWE, & DOEDENS, 2004; KIM & OSTERHOUT, 2005; VAN HERTEN, KOLK, & CHWILLA, 2005).

Recent event-related brain potentials (ERP)\(^7\) studies report that implausible verb-argument combinations, as in (9a-b), can deceive the processor and favor a semantically attractive interpretation (KIM & OSTERHOUT, 2005). This view predicts partially independent, but highly interactive syntactic and semantic processing streams. Under this assumption, in sentences such as (9), if syntactic cues fail to support a semantically attractive interpretation (‘surfer’ as the Agent of...

---


\(^7\) Event Related brain Potentials (GOMES, 2014) are a research methodology extracted from continuous EEG recordings in which stimulus presentation onset is coupled in time with brain responses so that the neurophysiological responses can be robustly related to stimuli. Segments are added and averaged per experimental condition so that they may be compared for condition effects. This is measured by comparing amplitudes (in voltage) and latencies (moment of maximum peak in ms) in a given time interval.
a 'shark-attacking’) and the syntactic cues are overwhelmed, the sentence would be perceived as syntactically ill-formed, although such eliciting conditions do not involve outright syntactic anomaly.

GOMES (2014) and GOMES & FRANÇA (2015) tested the chronology of the steps involved in the processing of role-reversed sentences. In this earlier study, we used event-related brain potentials (ERPs) in response to role-reversed sentences both in the active and in the passive voices to investigate whether there are distinct types of processing within sentence derivation. Comparing sentences, such as in (9a), *The surfer attacked the shark at the beach*, and in (9b) *The shark was attacked by the surfer at the beach*, respectively, to the canonical active ‘*The shark attacked the surfer at the beach*’ and passive ‘*The surfer was attacked by the shark at the beach*’, we aimed to investigate the nature of the brain's response to implausible, but syntactic well-formed role-reversed sentences. The focus of those experiments was to discriminate two well-known ERP responses, an N400 and a P600, associated to both semantic and syntactic processes.

Studies about semantic violations elicit a negative wave that peaks around 400ms after word onset. Semantically anomalous words show an N400 effect (KUTAS; HILLYARD, 1980), as it is called, while syntactically anomalous words elicit a positive ERP that onsets around 500ms after the word appears. This wave is called P600 and can be elicited by syntactic violations (Hagoort, P.; Brown; Groothusen, 1993). The neurophysiological responses reflecting syntactic and semantic processes seemed to be quite clear and simple, N400 effects are believed to be elicited due to semantic incongruences and P600 effects, due to syntactic ones. However, some very recent ERP studies seemed to contradict this general consensus. In light of these studies, Gomes (2014) and Gomes & França (2015) aim to contribute to the investigation of whether the N400 and the P600 responses are modulated by thematic role reversals, analyzing electrocortical manifestations of the time course of syntactically and semantically well-formed sentences in two experiments. For the sake of our argument here, we are only taking into analysis the results from the Active Voice Experiment.

Knowing the nature of the brain's response to implausible, but syntactic well-formed role-reversed sentences can give us a better understanding of the semantic and syntactic computation; it can also provide insight into the functional significance of the brain response that is elicited. For example, if an implausible, but syntactic well-formed sentence, such as, ‘*The surfer attacked the shark*’ were to elicit an N400 effect (KUTAS & HILLYARD, 1980), this would provide evidence
that the N400 is sensitive to lexico-semantic, morphosyntactic, pragmatic, and world knowledge information (e.g. VAN BERKUM, HAGOORT & BROWN, 1999; FEDERMEIER & KUTAS, 1999; HAGOORT, HALD, BASTIAANSEN & PETERSSON, 2004, FRISCH & SCHLESEWSKY, 2001, 2005; BORNKESSEL, MCELREE, SCHLESEWSKY, & FRIEDERICI, 2004). On the other hand, if a sentence like ‘The surfer attacked the shark’ were to elicit a P600 effect, it would provide evidence that either the syntactic processor is overwhelmed or it is under some reanalysis, as the sentence is syntactically well-formed.

GOMES (2014) and GOMES & FRANÇA (2015) found that infelicitous merge of the VP, such as [attacked the shark] yields an N400, and that the difficulty in making the integration between the stored external argument [the surfer] and the VP complex [attacked the shark], yields a P600. Hence, the authors argue in favor of a syntax-first account during which just lexical access of the subject and its partial integration with the verb would happen prior to the integration of the object. So the subject appears to be integrated in a piecemeal fashion, a process that only finishes completely after the integration of the object.

3. CURRENT EXPERIMENTAL STUDY

In light of the results found in GOMES (2014) and GOMES & FRANÇA (2015), here we want to add another layer of analysis in the study of subject-object asymmetry field by using cleft sentences that may advance the object linearly.

We intended to contribute to the discussion with different theoretical hypotheses on the analysis of cleft structures in different linguistic systems, based on empirical evidence of standard Brazilian Portuguese (BP).

In this study, we used the electrophysiological technique (EEG/ERP) to map distinguished syntactic and semantic processes, for instance, the N400 and the P600, addressing the focus as embodied in *It*-clefts clauses in three different conditions.

The stimuli exhibited Object *It*-Cleft clauses, as in the Table 1 below, with particular emphasis on (in) appropriate thematic assignment (semantic information) due to reversibility of canonical word order position (syntactic position).
I. Congruous Cleft Condition

“It was the SURFER that the shark attacked in Hawaii”

(PORT) Foi o SURFISTA que o tubarão atacou no Havaí.

II. Reversed Cleft Condition

“It was the SHARK that the surfer attacked in Hawaii”

(PORT) Foi o TUBARÃO que o surfista atacou no Havaí.

III. Incongruous Cleft Condition

“It was the COUCH that the shark attacked in Hawaii”

(PORT) Foi o SOFA que o tubarão atacou no Havaí.

For instance, sentences in condition I - Congruous Object Cleft Condition, they are considered Object *It*-Cleft sentences, since in these cases, it is the object of the relative clause (that the shark attacked [which]), that is in the focus position of the cleft clause (*it is X*).

The experimental hypotheses are exemplified by the reversed sentences in the condition II - Reversed Object Cleft Condition.

The III condition, Incongruous Object Cleft Condition is composed by incongruous object cleft sentences.

In summary, if information structural constraints are similar to the kind of semantic constraints that are known to modulate N400 amplitude or if processing difficulty results from a disruption of thematic role assignment — another semantically driven factor — we might expect to see a larger N400 response to information structure violations, (COWLES, ET AL, 2007; BORNKESSEL, SCHLESEWSKY, & FRIEDERICI, 2003; LI, HAGOORT, & YANG, 2008). On the other hand, if violations of information structural constraints disrupt larger processes of syntactic integration, then a modulation of components associated with such integration like the P600 could be seen, either immediately at the clefted noun or at the end of the sentence (ERICKSON & MATTSON, 1981, BREDART & MODOLO, 1988, BREDART & DOCQUIER, 1989, KUPERBERG, SITNIKOVA, CAPLAN, & HOLCOMB, 2003; HOEKS, STOWE, & DOEDENS, 2004). A P600 could also result if processing difficulty is caused by the reversal of plausibility relations in the information structure violations (KIM & OSTERHOUT, 2005, VAN HERTEN, KOLK, & CHWILLA (2005), KOLK, CHWILLA, VAN HERTEN, & OOR, 2003), used in our experiment.
4. MATERIALS AND METHODS

We recorded ERPs to critical segments while participants read It-Clefted sentences in three different conditions. We were concerned with the modulation of two distinct ERP effects, the N400 and the P600, both associated with semantic and syntactic processes, respectively. The experiments will be summarized next.

Our results should contribute to this discussion by deepening the understanding about the semantic and syntactic processes involved in role-reversed It-Cleft sentence processing.

4.1 The Experiment (Object It-Cleft).

We contrasted three groups of Brazilian Portuguese Object It-Clefted sentences: (I) Grammatically well-formed sentences (Foi o surfista que o tubarão atacou t1 no Havaí); (II) Role Reversed (Foi o tubarão que o surfista atacou t1 no Havaí); (III) Incongruous Role Reversed (Foi o sofá que o tubarão atacou t1 no Havaí).

The independent variables in this study are the thematic roles and the plausibility, while the subject response times, error rates and the ERPs are the dependent variables.

4.2.1 Participants

Prior to any testing, we approved our research plan with the local Ethics Committee in Brazil. This study was applied to 47 right-handed participants, native speakers of Brazilian Portuguese. Selection criteria also require all participants to have normal or corrected-to-normal vision. A written consent form was obtained from all subjects before participation.

4.2.2 Stimuli

The stimulus set of the Experiment (Object It-Cleft) was divided into three conditions, as demonstrated in Table 1. The agent-theme relation was controlled by frequency and cloze probability. Both of the arguments were controlled for animacy. Within each item in the canonical and the role-reversed sentences had an identical verb-argument and differed only in the order of such arguments. Further, in order to avoid wrap-up effects, the sentences were extended beyond
the post-verbal argument with an adverbial expression that was held constant across conditions within each set. The incongruous sentences carried an incongruous internal argument. Participants were tested with one of the three lists.

The trigger was set right at the onset of the It-Cleft Clause (i.e., Foi o surfista).

A total of 75 experimental sentences were adapted from the previous experiments in Brazilian Portuguese and 140 fillers were used across three lists. Each list will have 75 experimental items of the three different conditions exemplified in Table 1, and 140 fillers of two types: SVO word order and OVS word order.

Stimuli were pseudo-randomly ordered, subject to the following constraints: (1) experimental stimuli will be separated by at least one filler; (2) each third of the list presented from 10 to 12 targets of each condition. This is done so that there is no accumulation of experimental stimuli in any part of the test. Thus, by eliminating regularities, participants should not, implicitly or explicitly, get biased towards one particular manipulation.

### 5. RESULTS

Participants reliably judged canonical sentences to be plausible and the role-reversed sentences to be implausible with an overall accuracy of 90.9% (Canonical = 96.7%; Role-reversed = 88.5%). The ERPs for the sentences used in this Experiment are shown in Figure 1. In this experiment, results were analyzed with both participants (F1) and items (F2) as random effects. Two-way ANOVAS (3 segments x 3 conditions) revealed a significant main effect of segments and condition (F(3,207) = 8.04 p<0.000043 SS=2435634.12 MSe=100954.58; Condition F(3,207) = 0.591 p<0.621699 SS=554367.23 MSe=312792.23) with canonical sentences rated higher than role reversed sentences). The main effect of incongruousness was also significant (F(9,621) = 2.09 p<0.028723 SS=2392244.38 MSe=127352.77), reflecting the fact that congruous sentences were rated higher than incongruous ones. The interaction between segment and condition was also significant (both Fvalues < 0.05).
Regarding the role reversed condition, the results demonstrated that the reversals (as in “It was the shark that the surfer attacked in Hawaii”) impacted the amplitude of the N400 at the target verb (attacked) compared with the verb of the control condition (“It was the surfer that the shark attacked in Hawaii”) (F (1, 43) = 2.04  p< 0.0254). In the 600-800 ms interval, results did show an impact of the animacy-congruous role reversed arguments on the amplitude of the P600, at the verb, as shown in Figures 1 compared with the canonical condition (F (1, 43) = 4.76 p< 0.0001).

Additionally, the verb in the semantically anomalous condition (“It was the surfer that the shark attacked in Hawaii”) compared with the verb in the control condition elicited the canonical N400 effect at the verb (attacked), as shown in Figure 1 (F (1, 43) = )=4.15  p< 0.0001). In the 600-800 ms interval, results showed a visual effect at the verb, as shown in Figures 1 compared with the canonical condition, but it was not statistically relevant (F (1, 43) =0.60  p< 0.5529.

6. CONCLUSION

By means of an event-related brain potentials (ERP) study, we have proposed an experimental approach to test the subject-object asymmetry addressed by It-Cleft clauses, in Brazilian Portuguese.

Furthermore, we aimed at reestablishing a direct correspondence between syntax-first models and on-line sentence processing, relating them to new neurophysiological findings of two processing pathways, a ventral and a dorsal, the first for words and coordinated items and the second for hierarchical structures.

A number of studies have repeatedly observed an ERP component called the N400 for the lexical access and contextual integration of a filler into its gap (FEDERMEIER & KUTAS, 1999;
Juliana Novo Gomes e Aniela Improta França


Hence, we intended to relate the N400 and the P600 brain potentials (ERPs) to distinct types of processing within sentence derivation – semantic processes related to the thematic assignment and syntactic processes associated to sentence derivation.

Our hypotheses were be based on the assumption that, (1) If there was actually a difference of integration (i.e. gap-filling) costs between Subject It-Cleft clauses and Object ones, then we expected that asymmetries of processing load are included at the verb- position, which signals a cleft structure; and (2) If subject-object asymmetry was observed, then this asymmetry would reflect the difference in cost of integrating a filler into its gap.

We successfully found a N400 and a P600 at the verb (attacked) on the Reversed Condition compared with the Canonical one. We also found the canonical N400 at the verb, when comparing the Incongruous Condition with the Control one.

Taken together, these findings suggest that the presence of P600s to role-reversed sentences in previous studies could be attributed to the syntactic reanalysis, instead of the processing of the role reversals per se. Also, the presence of an N400 effect to the reversals could be due to the frustration of the strong combination of contextual constraints and strong lexical association. The results regarding the semantically anomalous sentences (“It was the surfer that the shark attacked in Hawaii”) are in accordance with the literature of N400 effects (see KUTAS & FEDERMEIER, 2011).

Our results make a unique contribution to the ERP response profiles, specially regarding the relationship between the role-reversals and the animacy violations in the Cleft Sentences. Our ERP findings seem to be compatible with the long-held assumption that the N400 and P600 appear to be modulated by the subject-object asymmetry, being sensitive to, respectively, the semantic attraction between words in the sentences and, the congruency of the predicate and both its arguments concatenation.
Therefore, our results are not compatible with the hypothesis that assumes an independent semantic composition, according to which readers can be initially blind to the implausibility in role-reversed sentences when they first encounter the verb. Instead, our findings support the assumption that semantic interpretation is based on the syntactic structure of the sentence. We thus claim that syntactic anomalies block the detection of semantic anomalies, therefore, semantically incongruous sentences, such as role-reversals are perceived to be odd due to a syntactic constraint satisfaction that assigns the right theta-roles to the verbs arguments despite the semantic cues, because, in fact, its syntax driven.

REFERENCES:


Casteleiro, João Malaca (1979) Sintaxe e Semântica das Construções Enfáticas com “é que”. Boletim de Filologia XXV.


Juliana Novo Gomes e Aniela Improta França

Processing it-cleft sentences in Brazilian Portuguese: an ERP study of leftward-moved constituents in role-reversed sentences


Robert, Reichle (2014) Cleft type and focus structure processing in French. Language, Cognition and Neuroscience, Vol. 29, Iss. 1, 2014


Stroud, C.; C Phillips, C. Examining the evidence for an independent semantic analyzer: An ERP study in Spanish. [s.l]: [s.n.].


