On the movement analysis of null subjects in Brazilian Portuguese: Experimental results from extraction of embedded subjects[*]

Claudia Coelho
Universidade de São Paulo
claudia.coielho@usp.br

Jairo Nunes
Universidade de São Paulo
jmnunes@usp.br

Leticia Santos
Universidade de São Paulo
leticia.evelyn.santos@usp.br

Abstract

There is an on-going debate on the empirical adequacy of the movement approach to definite null subjects in Brazilian Portuguese, *BP* (cf. Ferreira 2000, Rodrigues 2004, and Nunes 2009). On the one hand, Modesto (2000) observes that the null subject of finite embedded clauses associated with object control verbs like *convencer* 'to convince' is subject-oriented, rather than object-oriented and this is unexpected from a movement perspective. On the other hand, Rodrigues (2004) claims that these embedded clauses are actually adjuncts in BP and the observed subject orientation can be accounted for in terms of Hornstein's (1999, 2001) movement analysis of adjunct control. This paper aims to contribute to this debate by presenting the results of an experiment on grammaticality judgments by BP speakers on the extraction of embedded subjects out of complement and adjunct clauses, as well as finite embedded clauses associated with *convencer*. The results show that when a distinctive pattern could be observed, finite clauses associated with *convencer* behaved like adjunct clauses rather than complement clauses. The experiment thus provides confirming evidence for Rodrigues's (2004) adjunct analysis, invalidating Modesto's (2000) argument against the movement approach to definite null subjects in BP.
Keywords: null subjects; Brazilian Portuguese; Movement Theory of Control; extraction of embedded subjects; grammaticality judgment experiment

Table of Contents

1. Introduction

The considerably large literature on null subjects in Brazilian Portuguese (henceforth BP) converges on the conclusion that they do not behave like the null subjects of canonical pro-drop languages[2]. (1a) below, for instance, shows that a definite (i.e. nonexpletive, nonarbitrary) third person null subject in BP is banned from matrix environments that exclude topic drop. This amounts to saying that definite null subjects in BP typically appear in embedded clauses. (1b) further shows that the embedded null subject requires an antecedent and that this antecedent must be the most local c-commanding DP.

---

[2] For relevant discussion on the availability and interpretation of null subjects in BP, cf. Chao (1983), Moreira da Silva (1983), Negrão (1986), Galves (1987, 2001), Duarte (1995), Figueiredo Silva (1996), Kato (1999), Ferreira (2000, 2009), Kato & Negrão (2000), Modesto (2000, 2011), Barbosa, Duarte & Kato (2005), Rodrigues (2002, 2004), Nunes (2008, 2009, 2011), Holmberg, Nayudu & Sheehan (2009), Petersen (2011), and Saab (2016).

(1)   a. *O que Ø comprou ontem?
          what   bought   yesterday
          'What did she/he bought yesterday?'
       b. [O João]$_i$ disse   que   [o pai   d[o Pedro]$_j$]$_k$ acha que
          *the João said   that   the father of-the Pedro   thinks tha*t
          Ø$_k$/*$_i$/*$_j$/*$_l$ vai ser promovido.
                          goes be promoted
          'João said that [Pedro's father]$_k$ thinks that he$_k$ is going to be promoted'

Facts like the ones illustrated in (1) have led Kato (1999) to propose that null subject constructions in BP involve PRO rather than pro. Reinterpreting the gist of Kato's proposal in terms of Hornstein's (1999, 2001) Movement Theory of Control (henceforth MTC), Ferreira (2000, 2009) and Rodrigues (2002, 2004) have argued that definite null subjects in BP are traces of A-movement. Under this approach, contrasts such as the one illustrated in (2) below, for instance, are accounted for in terms of minimality. The null subject in both (2a) and (2b), being a trace of A-movement, requires that its antecedent be the closest c-commanding DP. In other words, the null subject must be interpreted as *a Maria* 'Maria' in (2a) and *o médico* 'the (male) doctor' in (2b). However, the lexical meaning of *grávida* 'pregnant' as well as its feminine agreement morphology are only compatible with the former; hence the contrast between (2a) and (2b).

(2)   a. O médico   disse que a   Maria acha   que *t* está grávida
          the doctor.MASC said that the Maria thinks that   is pregnant.FEM
          'The doctor said that Maria thinks that she is pregnant'
       b. *A   Maria disse que o   médico   acha   que *t* está grávida
          the   Maria said that the doctor.MASC thinks that   is   pregnant.FEM
          'Maria said that the doctor thinks that she is pregnant'

In this paper we examine an empirical challenge raised by Modesto (2000, 2011) to approaches that analyze null subjects in BP as traces of A-movement (cf. Ferreira 2000, 2009, Rodrigues 2002, 2004, and Nunes 2008, 2009). The challenge is based on his empirical observation that the null subject of finite embedded clauses associated with object control verbs like *convencer* 'to convince' are subject oriented, rather than object oriented. Assuming that the matrix object of these constructions c-commands the embedded clause, Modesto concludes that it should block movement of the embedded subject to the matrix subject position; in other words, the subject orientation reading should be prohibited, contrary to fact. Rodrigues (2004) objects to this reasoning, claiming that these embedded clauses are actually adjuncts in BP and that the observed subject orientation can be accounted for in terms of Hornstein's (1999, 2001) movement analysis of adjunct control. Extraction of embedded subjects in these constructions should in principle suffice to settle this debate, but the reported judgments in this regard are conflicting (fully acceptable according to Modesto 2011 and marked as ?? according to Rodrigues 2004). Since there is no agreement on judgments regarding crucial pieces of data, we designed a grammaticality judgment experiment with the goal of shedding more light on the discussion. The results to be presented below show that when a distinctive pattern could be observed, finite clauses associated with *convencer* behaved like adjunct clauses rather than complement clauses.

The paper is organized as follows. In section 2, we present the details of Modesto's (2000) challenge to the movement approach to null subjects in BP and Rodrigues's (2004) alternative analysis of the apparently problematic data. We show that A'-extraction of embedded subjects should in principle tease the two approaches apart, but a proper assessment of its relevance becomes clouded as the judgments presented by these authors go in opposite directions. In section 3, we then present the experiment on judgments on such extractions that we carried out. In sections 4 and 5 we present and discuss the results of the experiment. Finally, section 6 concludes the paper.

2. A challenge to the movement analysis of null subjects in BP

Hornstein (1999, 2001) has argued within the MTC that object control follows from minimality. Under a Larsonian shell analysis of ditransitive structures, the matrix object of a sentence like (3a), for instance, should c-command into the infinitival clause, as sketched in (3b). If obligatorily controlled PRO is a trace of A-movement, as defended by Hornstein, the subject position of the infinitival clause in (3b) must be occupied by the trace of *a Maria*, for movement of *o Paulo* across *a Maria* should violate minimality. Hence, structures like (3) give rise to object control rather than subject control.

(3)   a. [O Paulo]$_1$ convenceu [a Maria]$_2$ a $\textbf{\textit{ec}}_{2/*1}$ sair
             the Paulo   convinced the Maria to   leave
            'Paulo convinced Maria to leave'
         b. [$_{vP}$[o Paulo]$_1$ convenceu-v [$_{VP}$ [a Maria]$_2$ [$_{V'}$ $t_{convenceu}$ [$t_{2/*1}$ a sair]]]]

Modesto (2000: 20) takes the contrast between (3a) and (4) below to provide evidence against the movement approach to null subjects in BP finite clauses outlined in (2) (cf. Ferreira 2000, 2009, Rodrigues 2002, 2004, Nunes 2008, 2009). He points out that if the embedded null subject of (4) is an A-trace, it should also find its antecedent in the matrix object position, contrary to fact. His conclusion is that the null subject of the infinitival clause in (3a) may be prone to a movement analysis, but not the null subject of the finite clause in (4).

(4)   [O Paulo]$_1$ convenceu [a Maria]$_2$ que $\textbf{\textit{ec}}_{1/*2}$ tinha que ir embora
          the Paulo   convinced the Maria   that   had   that go away
         'Paulo convinced Maria that he had to go away'

Modesto further makes the important observation that if the matrix object of sentences such as (4) undergoes A'-movement, it may then be interpreted as the antecedent for the embedded null subject, as shown in (5).

(5)   Quem$_1$ que [o Pedro]$_2$ convenceu t_1 que $\textbf{\textit{ec}}_{1/*?2}$ tinha que ir embora?
          who   that the Pedro convinced that   had   that go away
         'Who did Pedro convince that he had to go away?'

                                                                    (Modesto 2000:85)

Modesto takes the contrast between (4) and (5) as evidence for his claim that null subjects in BP are pros that must be licensed via A'-binding. He proposes that subjects of finite clauses in BP may move to Spec,AgrP, which is taken to be

an A'-position, and from this position they can license an embedded null subject. That would be the case of (4). By contrast, in (5) the matrix object passes through Spec,AgrP on its way to Spec,CP and the matrix subject stays put in its Case checking position; hence, only the matrix object should be able to function as an antecedent for the embedded null subject.

Before we discuss the contrast between (3a) and (4), it should be observed that for purposes of exposition, we have maintained Modesto's original judgments in (5), according to which only the object reading is licit when the object undergoes A'-movement. Although there may be a preference for the object reading for some speakers, Rodrigues (2004: 217) notes that the sentence in (6) below[3] clearly shows that the subject reading is available, as well.

(6)   Quem$_i$ [**a Maria**]$_k$ convenceu t_i que **ec**$_k$ estava grávida?
         who the Maria   convinced that   was   pregnant
        'Who did Maria convince that she was pregnant?'

Going back to the contrast between (3a) and (4), the tacit assumption in Modesto's argument is that they have the same structural configuration (see (3b)), differing only in terms of finiteness. There are reasons to believe that this is not the case, though. Ferreira (2000) shows that the matrix object does not c-command into the finite embedded clause, as indicated by the lack of a Principle C effect in a sentence such as (7) below. Besides, Rodrigues (2004) shows that the embedded clause of (4) does not behave as a typical complement, for it does not allow extraction from within it, as shown in (8) (cf. Nunes 2009 for further discussion)[4]. Accordingly, we find a clear contrast between (7)-(8) and sentences with infinitivals like (3a) (cf. Nunes 2013). As (9) shows, in the case of infinitivals, the matrix object c-commands into the infinitival clause, inducing a Principle C effect in (9a), and the embedded clause is transparent for extraction (see (9b)).

(7)   O João convenceu [**a Maria**]$_i$ [que [**a idiota**]$_i$ deveria assaltar um banco]
        the João convinced the Maria   that the idiot should rob   a   bank
        'João convinced Maria that the idiot should rob a bank'
                                                          (adapted from Ferreira 2000: 39)
(8)   ??**Quem**$_i$ o   João convenceu a Maria [que t_i vai viajar]?
         who   the João convinced the Maria that goes travel
          'Who did João convince Maria that will travel?'
                                                          (Rodrigues 2004: 219)

---

[3] Incidentally, observe that (6) sharply contrasts with (2b). Thus, one could not simply attribute the acceptability of (6) to the pragmatic compatibility between the embedded predicate *grávida* 'pregnant' and the DP *a Maria*; otherwise, (2b) should also be acceptable, contrary to fact.
[4] As show in (9b), BP does not exhibit *that*-trace effects. Hence, the ungrammaticality of (8) must be due to something else (see below).

(9)   a. *O João convenceu **[a Maria]**$_k$ [*ec* a espalhar que **[a idiota]**$_k$ ia
           the João convinced the Maria   to spread   that the idiot went
        renunciar
        resign
        'João convinced Maria to spread the news that the idiot was going to
         resign.'
     b. **Quem**$_i$ o   João convenceu a   Maria [*ec* a dizer que *t*$_i$ era o
        who   the João convinced the Maria   to say that   was the
        melhor   candidato]?
        best   candidate
        'Who did João convince Maria to say was the best candidate?'

Even more interesting, Rodrigues (2004) shows that the additional object reading available in sentences like (5) when the object undergoes A'-movement is also observed in BP with respect to adjunct finite clauses. The pair of sentences in (10), for instance, shows that the null subject of the finite adjunct clause obligatorily takes the matrix subject as its antecedent if the matrix object remains *in situ* (see (10a)), but may be interpreted as either the matrix subject or the matrix object if the latter undergoes A'-movement (see (10b)).

(10)   a. [A Maria]$_i$ visitou quem$_k$ [quando **ec**$_i$/*$_k$ foi   para Brasília]?
            the Maria visited who   when   went to   Brasília
           'Who did Maria visit when she went to Brasília?'
       b. Quem$_k$ [a   Maria]$_i$ visitou *t*$_k$ [quando **ec**$_i$/$_k$ foi   para Brasília]?
          who   the Maria visited   when   went to   Brasília
          'Which person did Maria visit when she/that person went to Brasília?'
                                                          (Rodrigues 2004: 228)

In the face of these facts, Rodrigues (2004) proposes that the finite embedded clause associated with verbs like *convencer* 'to convince' in BP is an adjunct of sorts[5]. This would account for why the object does not induce a Principle C effect in sentences such as (7) and why extraction out of the embedded clause as in (8) does not yield acceptable outcomes. As for the additional reading in (5), it should in principle be subject to whatever explanation is offered for the pattern in (10b) (cf. Rodrigues 2004 and Nunes 2013, 2014, 2016 for specific proposals).

Importantly, Rodrigues shows that if the finite embedded clause of sentences like (4) is an adjunct, its subject orientation can be properly accounted for under the movement approach. Hornstein (1999, 2001) has argued that adjunct

---

[5] Nunes (2009) has argued that the finite embedded clause associated with *convencer* may display a complement behavior when preceded by the preposition *de* 'of'. In particular, embedded subject extraction is allowed, as illustrated in (i) below. In this paper, we will restrict our discussion to constructions without *de*. See Coelho (in progress) for a grammaticality judgment experiment testing the role of *de* in these constructions.

(i)   Quem$_i$ o João convenceu a   Maria **de** [que *t*$_i$ vem   amanhã]?(Nunes 2009: 257)
        who   the J. convinced the M.   **of** that comes tomorrow
        'Who did João convince Maria will come tomorrow?'

control is to be derived in terms of sideward movement (in the sense of Nunes 2001, 2004) and that subject orientation in adjunct control follows from Merge-over-Move economy computations (cf. Chomsky 1995). Extending Hornstein's adjunct control analysis to finite adjunct control in BP, Ferreira (2000) and Rodrigues (2004) argue that the null subject of a finite adjunct such as the one in (10a) is a trace of the matrix subject, which has undergone sideward movement *before* the embedded clause becomes an adjunct. From this perspective, the derivation of (10a) proceeds along the (simplified) lines of (11).

(11)   a. N = {**quem**$_1$, ...}
                who
             K = visitou   L = [[**a   Maria**] foi   para Brasília]
                 visited   the Maria   went to   Brasília

          b. N' = {**quem**$_0$, ...}
             K' = [visitou **quem**]   L = [[a Maria] foi para Brasília]

          c. K'' = [[**a Maria**]$_i$ visitou quem]   L = [t_i foi para Brasília]

          d. [$_{TP}$ [**a   Maria**]$_i$ [$_{vP}$ [$_{vP}$ t_i visitou quem][quando t_i foi   para Brasília]]]
                the Maria   visited who   when   went to   Brasília

        Given the derivational step in (11a), the verb *visitou* can have its selectional requirements satisfied either via selection and merger of *quem* or via sideward movement of a Maria. Assuming that Merge is more economical than Move (cf. Chomsky 1995), the computational system chooses merger of *quem* (see (11b)) before *a Maria* undergoes sideward movement to external argument position of *visitou* (see (11c)). Notice that movement of *a Maria* in (11c) crucially takes place while L is a root syntactic object – in other words, before L becomes an island (cf. Hornstein 2001 for detailed discussion). Finally, *a Maria* moves to the matrix Spec,TP, yielding the subject reading for the embedded subject.
        Applying this analysis to constructions involving finite clauses associated with verbs like *convencer*, Rodrigues argues that the subject orientation observed in (4), for instance, is the outcome of a derivation along the lines sketched in (12) below. Given the derivational step in (12a), Merge-over-Move triggers merger of *a Maria* before *o Paulo* undergoes sideward movement to the external argument position of *convencer*. Thus, the embedded null subject of (4) is interpreted as the matrix subject for it is a trace of the matrix subject[6].

---

[6] Crucially, a derivation along the lines of (12) is not available for a sentence involving a true complement clause. Take (2b), repeated here in (ia), for instance. As shown in (ib), the most embedded clause is merged as the complement *acha* and the intermediate clause, as the complement of *disse*. At this derivational step, *a Maria* cannot move to the matrix [Spec,*v*P] due to the intervention of *o médico* in the intermediate clause; hence the ungrammaticality of (ia).

(i)   a. *A Maria disse que o   médico   acha   que *t* está grávida
            the Maria said that the doctor.MASC thinks that   is   pregnant.FEM
            'Maria said that the doctor thinks that she is pregnant'
          b. [$_{vP}$ *v* [$_{VP}$ disse [$_{CP}$ que [o médico] acha [$_{CP}$ que [a Maria] está grávida]]]

(12)   a. N = {**a₁, Maria₁,** ...}
         the Maria
      K = convenceu   L = [que [o Paulo] tinha que ir embora]
        convinced   that the Paulo had that go away

    b. N' = {**a₀, Maria₀,** ...}
      K' = [convenceu **[a Maria]**]   L = [que [o Paulo] tinha que ir embora]

    c. K'' = [[**o Paulo**]ᵢ convenceu [a Maria]] L = [que t_i tinha que ir embora]

    d. [$_{TP}$ [**o Paulo**]ᵢ [$_{vP}$ [$_{vP}$ t_i convenceu [a Maria]] [que t_i tinha que ir embora]]]

In a later paper, Modesto (2011) questions Rodrigues's (2004) judgments on subject extraction in sentences like (8) (which are the same as ours). According to him, the sentence (13a), which is parallel to (8), is "perfectly grammatical", whereas the sentence in (13b), which involves subject extraction out of a *bona fide* adjunct, is "grossly ungrammatical" (p. 16).

(13)   a. Quem₁ (que) o   Pedro convenceu a   Cilene que t_1 vai   viajar?
      who   that the Pedro convinced the Cilene that   goes travel
     'Who is the person that Pedro convinced Cilene that he will travel?'
    b. *Quem₁ o   Pedro viu a   Cilene enquanto t_1 fazia compras?
      who   the Pedro saw the Cilene while   did   shopping
     *'Who did Pedro see Cilene while was shopping?'
                                  (Modesto 2011: 15-16)

Establishing the grammaticality status of sentences such as (8) and (13a) therefore becomes a central issue on the debate on grammatical nature of null subjects in BP. Bearing that in mind, in this paper we subject the contradicting judgments reported by Rodrigues (2004) and Modesto (2011) to a closer scrutiny. We developed an acceptability experiment to test judgments by BP speakers on extraction of subjects of finite clauses, comparing their judgments on standard complement and adjunct clauses with their judgments on parallel cases involving *convencer*[7]. The prediction is clear: extraction out of cases involving *convencer* should pattern like extraction out of complements if Modesto's (2000) analysis of these cases is correct, but like extraction out of adjuncts if Rodrigues's (2004) analysis is correct. In the next sections, we describe the experiment and present the results found.

---

[7] Modesto (2000, 2011) notes that verbs like *avisar* 'to warn', *informar* 'to inform', *alertar* 'to alert', and *prevenir* 'to forewarn', among others, exhibit the same pattern as *convencer* 'to convince'. However, given that his argument is formulated based on examples exclusively involving *convencer,* we decided to also use only this verb in the test cases in order to avoid a potential noise introduced by eventual lexical differences within this class of verbs.

3. The experiment

In order to investigate the judgments on the types of sentences discussed above, we designed an acceptability test and applied it to 20 participants (10 males and 10 females) with 20 to 30 years of age, all native speakers of BP who were born and raised in the state of São Paulo[8]. The participants were told that we were evaluating the performance of a computer software designed to formulate complex questions in Portuguese and in order to do that, we needed Portuguese speakers to judge which questions built up by the software were well formed and which were not. The participants listened to audio recorded sentences read by a BP speaker and pressed YES or NO in the computer to indicate if the sentence listened to was well formed or not. The participants could also press REPETE in case they wanted to listen to the sentence once again before pressing YES or NO.

Before the presentation of the relevant sentences of the experiment, there was a familiarization period where the participants were presented with 8 questions unrelated to our research topic (4 with expected YES-answers and 4 with expected NO-answers). This allowed us to check if the participants had correctly understood the instructions of the experiment.

The actual experiment consisted of 10 sentences with subject extraction out of a finite complement clause, 10 sentences with subject extraction out of a finite adjunct clause, 10 sentences with subject extraction out of a finite clause associated with the verb *convencer*, and 8 distractors (4 with expected YES answers and 4 with expected NO answers). (14) below shows an example of each type of sentence. All the sentences with complement clauses involved a ditransitive verb (see (14c)) and all the sentences with adjunct clauses involved a transitive verb in the matrix clause (see (14d)). This ensured that an eventual distinct behavior of sentences with *convencer* should not be attributed to the complexity induced by the presence of an object in the matrix clause. We also computed if the extracted *wh*-subject was syntactically simple (*quem* 'who' or o que 'what') or complex (*que pessoa* 'which person' or *que prato* 'which dish', for instance). The sentences of the different paradigms were mixed and all participants were exposed to the same order of presentation of the sentences.

(14)   a. *Familiarization sentence with an expected YES answer*:
       Que animal a   Maria viu quando foi   ao   zoológico com o   Pedro?
       which animal the Maria saw when   went to-the zoo   with the Pedro
      'Which animal did Maria see when she went to the zoo with Pedro?'

    b. Familiarization sentence with an expected NO answer:
     *O que o   João convidou a   Maria pra comer o   bolo?
       what   the João invited   the Maria to   eat   the cake
     *'What did João invite Maria to eat the cake?'

---

[8] The experiment was approved by the Ethics Committee of the Escola de Artes, Ciências e Humanidades da Universidade de São Paulo under protocol CAAE: 68901117.9.0000.5390.

c. *Sentence with subject extraction out of a complement clause:*
    Quem a   Maria contou pro   Pedro que vem   pra   festa de amanhã?
    who   the Maria told   to-the   Pedro that comes to-the party of tomorrow
    'Who$_i$ did Maria tell Pedro t_i will come to the party?'

d. *Sentence with subject extraction out of an adjunct clause:*
    *Quem a   Maria encontrou o João depois que   viajou pra Nova York?
     who   the Maria met   the João after   that traveled to New York
    *'Who$_i$ did Maria meet João after t_i traveled to New York?'

e. *Sentence with subject extraction out of an embedded clause associated*
   with convencer:
    Quem o   João convenceu a   Maria que vai   viajar   pra   Holanda?
    who   the João convinced the Maria that goes travel to-the Netherlands
   'Who$_i$ did João convince Maria t_i will travel to Netherlands?'

f. *Distractor with an expected YES answer*:
    Quem a   Maria disse que o   João chamou pra   sair?
    who the Maria said that the João called   to   leave
    'Who Maria said that João asked to go out?'

g. *Distractor with an expected NO answer*:
    *Quem a Maria terminou com o   Pedro depois que ela conheceu na
     who   the Maria finished with the Pedro after that   she met   in-the
     festa?
     party
    *'Who$_i$ did Maria break up with Pedro after she met t_i at the party?'

There was a total of 920 answers. The data were collected and categorized with the free software TP from Worken (http://www.worken.com.br/tp_regfree.php), developed for speech perception experiments.

4. Results

4.1. The actual sample for the analysis
The answers of one participant were excluded, for she pressed YES for all the sentences, which suggests that she did not understand her task. Out of the set of 46 sentences, the three sentences in (15) below were also excluded, for they exhibited an atypical behavior with respect to both the global picture (see Figure 1) and the individual participants (see Graphic 1)[9].

---

[9] Coincidentally, the outliers in (15) all involve sentences with *convencer*. The elimination of these three sentences did not affect the overall results, though, as we can see in Figures I and II below.

(15) **a. s21**:   Que   funcionário$_i$ o   João convenceu a Maria
which   employee   the   João convinced the Maria
[que t_i devia   sair   da   sala]?
that   should leave   of-the room
'[Which employee]$_i$ did João convince Maria t_i should leave the
room?'

   b. **s31**:   Que   político$_i$ o   João convenceu   a   Maria
which   politician the   João convinced   the Maria
[que t_i tem   ficha   suja]?
that   has   file   dirty
'[Which politician]$_i$ did João convince Maria t_i has a criminal
record?'

   c. **s32**:   Que   aluno$_i$ a   professora   convenceu a   diretora
which   student the   professor   convinced the director
[que t_i fez   uma   boa   apresentação]?
that   made a   good presentation
'[Which student]$_i$ did the teacher convince the director t_i made a
good presentation?'

Figure 1. Boxplot for all the sentences
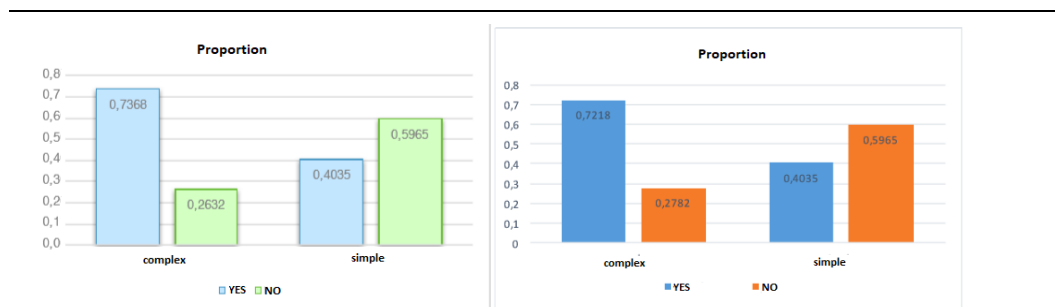
Figure I. Overall proportion of YES-

and NO-answers with all the sentences

Figure II. Overall proportion of YES-

and NO-answers without outliers
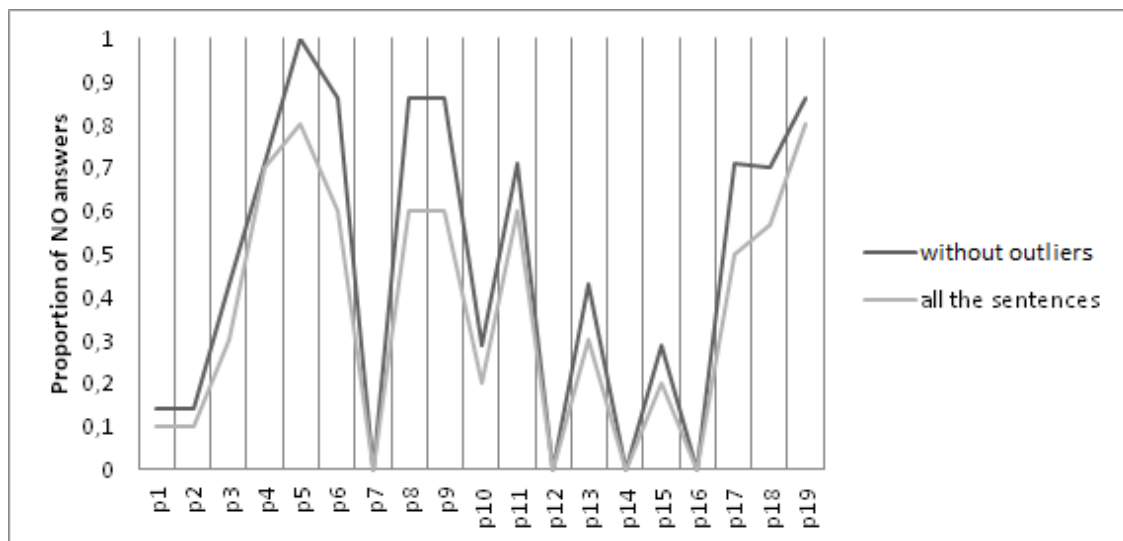
Graphic 1. Proportion of NO-answers with respect to *convencer*-sentences by individual participants

With the exclusion of the participant mentioned above and the outliers sentences in (15), the actual sample to be examined consisted of 817 answers (43 sentences times 19 participants), distributed between simple and complex *wh*-phrases as depicted in Table 1

Distribution of complex and simplex *wh*-phrases			
Sentence type	Simplex *wh*-phrases	Complex *wh*-phrases	Total
familiarization	6	2	8
complement	5	5	10
adjunct	6	4	10
convencer	3	4	7
distractor	3	5	8
Total	23	20	43

Table 1. Distribution of complex and simplex *wh*-phrases across sentence types

4.2. Homogeneity and reliability of the data
The remaining 19 participants behaved homogeneously and no participant was identified as an outlier, as we can see in Figure 2.
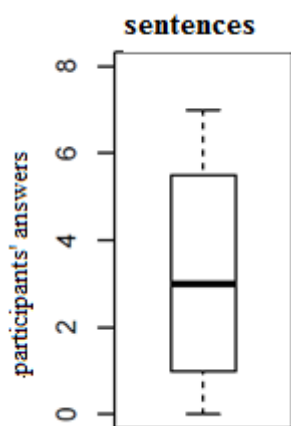
Figure 2. Boxplot for the participants

Within each type (familiarization, distractor, complement, adjunct, and *convencer*), the individual sentences also displayed a homogeneous behavior, as respectively shown in Figures 3-7.
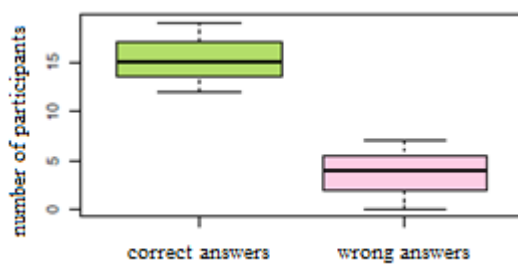
Figure 3. Boxplot for familiarization sentences
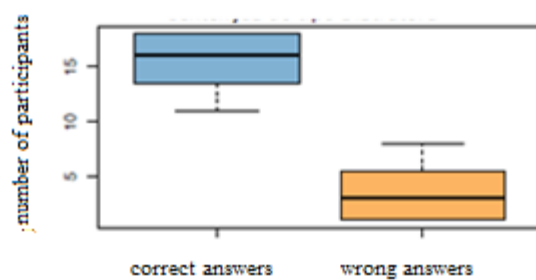
Figure 4. Boxplot for distractor sentences
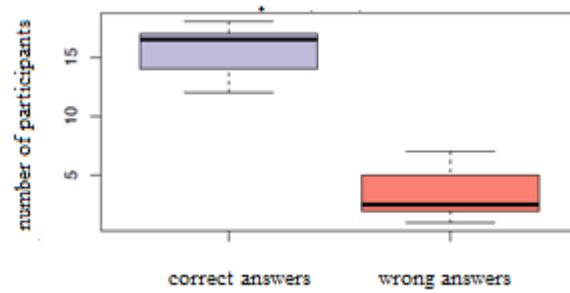
Figure 5. Boxplot for complement sentences
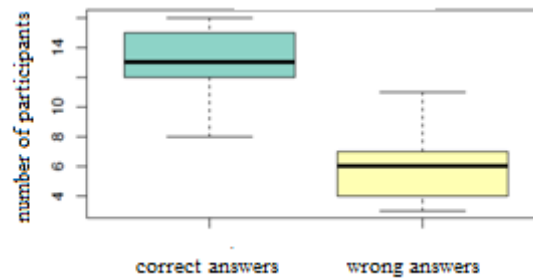
Figure 6. Boxplot for adjunct sentences
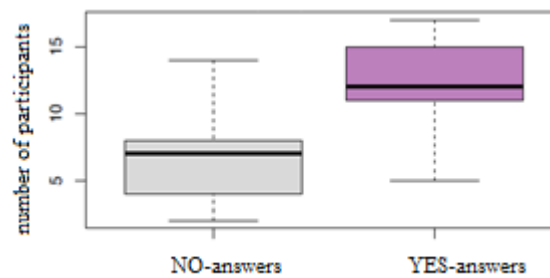
Figure 7. Boxplot for sentences with *convencer*

Figure 8 below displays the proportion between right and wrong answers with respect to familiarization and distractor sentences (0.8/0.2 and 0.82/0.18, respectively). These results indicate that the participants understood their task and were paying attention, which in turn suggests that their answers to the other sets of sentences are reliable.
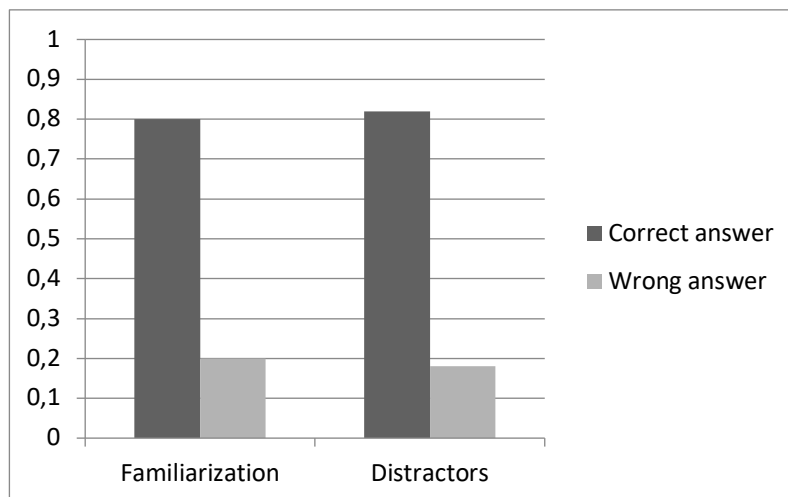
Figure 8. Proportion between right and wrong answers for familiarization and distractor sentences

The dispersion Graphics 2 and 3 further show that there is no correlation between the correct answers provided by individual participants with respect to familiarization and distractor sentences and their answers to *convencer*-sentences (the coefficient of linear correlation between x and y axes is 0,24 in the former and 0,11 in the latter). In other words, it is not the case that a given participant is providing more YES or NO answers to *convencer*-sentences because (s)he did not understand the task of the experiment or was not paying sufficient attention.
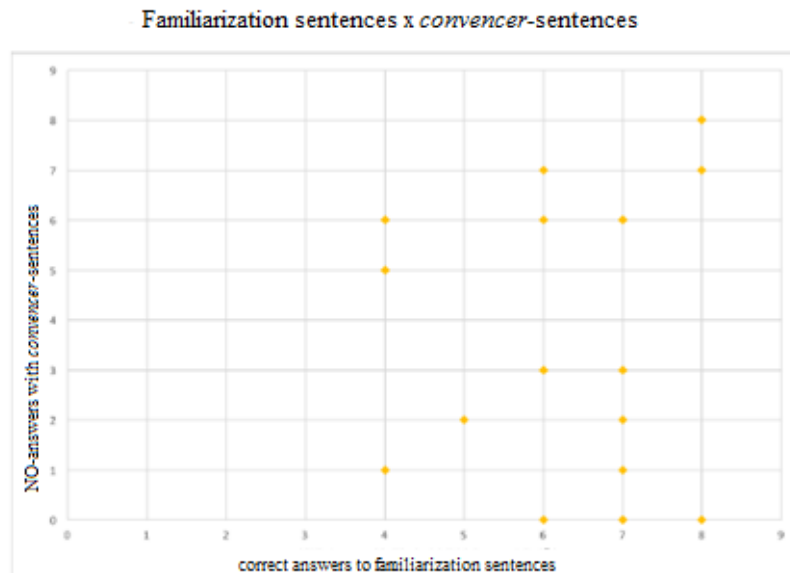
Graphic 2. Individual behavior with respect to familiarization and *convencer*-sentences

Distractor sentences x *convencer*-sentences
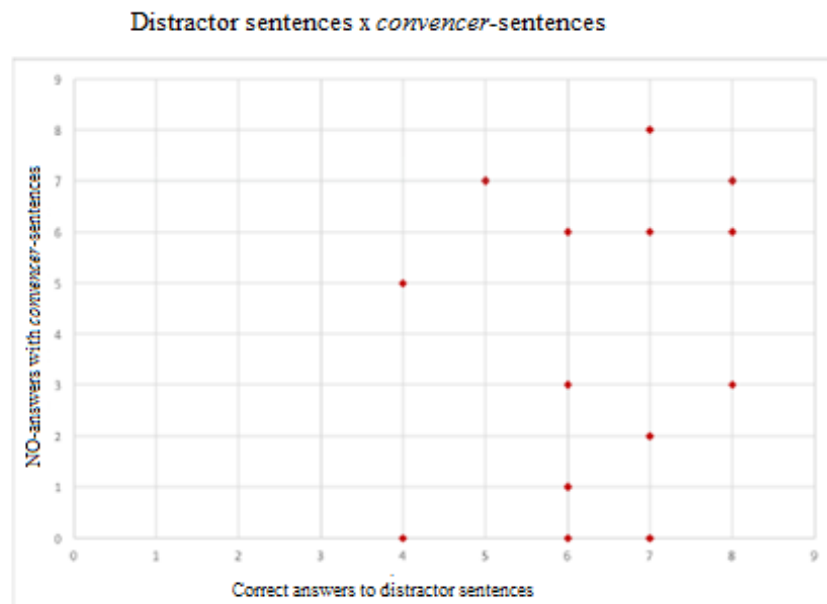
Graphic 3. Individual behavior with respect to distractor and *convencer*-sentences

4.3. Repetition and syntactic complexity of the extracted wh-*phrase*
We examined whether there was a correlation between the syntactic complexity of the *wh*-phrase and repetition. One could expect that simple *wh*-phrases should trigger repetition more often than complex *wh*-phrases, given that the latter is more prone to a D-linking interpretation (in the sense of Pesetsky 1987). This expectation was not fulfilled, though. Figure 9 shows that the type of *wh*-phrase is not a determining factor in triggering repetition (p = 0.406910).
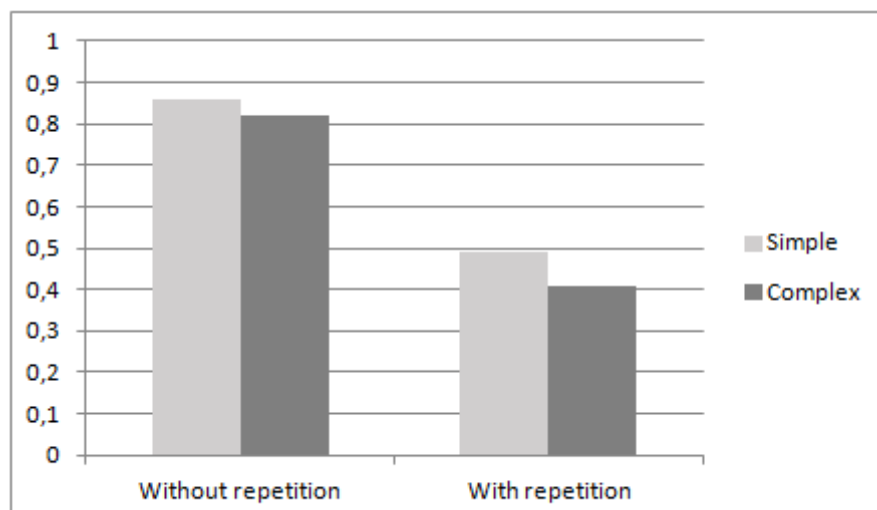
Figure 9. Proportion of repetition per type of *wh*-phrase

4.4. Repetition and sentence type

Figure 10 below shows that familiarization sentences are the ones that display the larger proportion of repetition (0.48/0.52). This is an expected result, as it indicates that the participants wanted to make sure they got familiarized with the type of sentences they were supposed to evaluate. This is further confirmed in Graphic 4, which shows that the number of repetitions decreases as the experiment unfolds[10].
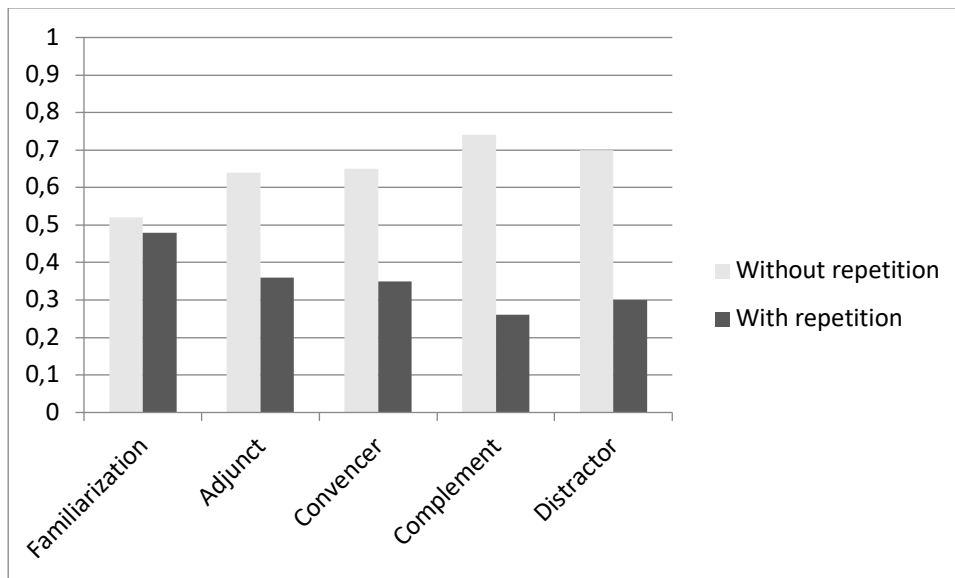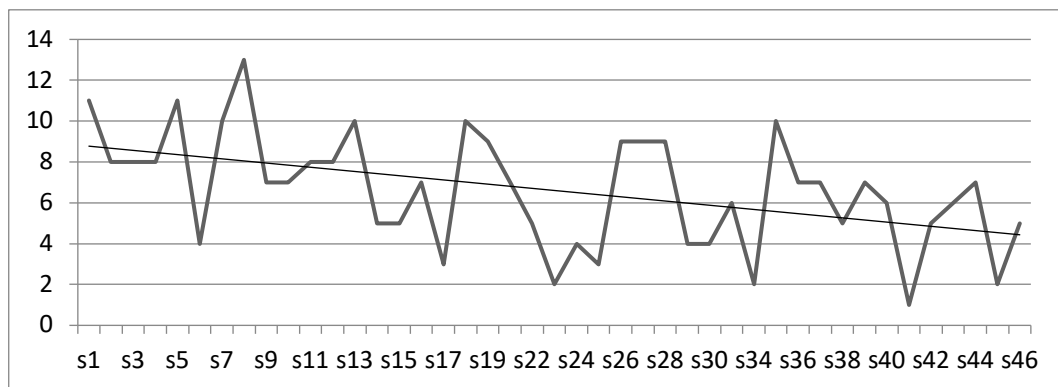

Figure 10. Proportion of repetition by sentence type

Graphic 4. Repetition along the experiment

Recall that only the familiarization sentences were presented together at the beginning of the experiment (the first 8 sentences); the sentences of the other paradigms were intermingled (and presented in the same order to all participants).

---

[10] Recall (see section 4.1) that out of the 46 original sentences, we have excluded three outliers (s21, s31, and s32). For easiness of reference, we maintained the original identification of the other sentences, though. So, *s46* in Graphic 4, for instance, actually refers to the 43rd sentence of the sample (the last one).

The directionality seen in Graphic 4 suggests that as the experiment unfolds, the participants become more and more secure with respect to the task itself, pressing REPETE only when in doubt with respect to specific sentences. In this regard, note that the proportion of repetition for *convencer*-sentences (0.35/0.65) in Figure 10 above is much more similar to the proportion for adjunct sentences (0.36/0.64; p = 0.726557) than the proportion for complement sentences (0.26/0.74; p = 0.119698). In fact, the proportion of repetition for complement sentences is much closer to the proportion for distractors (0.3/0.7; p = 0.418481).

4.5. Tendency for YES- and NO-answers during the experiment

We also examined whether the order of presentation of the sentences affected the participants' answers, that is, whether there was a tendency for participants to give more YES- or NO-answers for each type of sentence in the beginning of the experiment than at the end.

Graphics 5 and 6 below respectively depict the proportion of expected answers for complement and adjunct sentences as the experiment unfolds. As extraction is allowed out of complement clauses but not out of adjunct clauses, YES was the expected answer for complement sentences and NO for adjunct sentences. Given that the aim of the experiment was to determine the behavior of *convencer*-sentences, there was no a priori expected answer in this case. However, in order to make the appropriate comparison with complement and adjunct sentences, it was necessary to arbitrarily choose between YES and NO. For concreteness, we computed NO as the expected answer for *convencer*-sentences and the result is shown in Graphic 7 below.
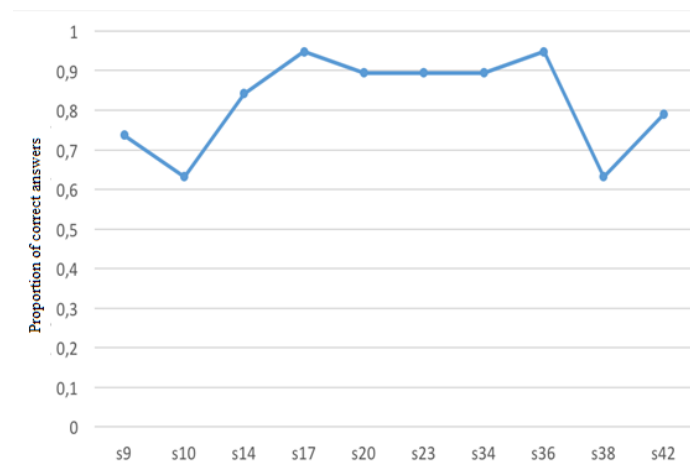
Graphic 5. Proportion of expected answers for complement sentences along the experiment
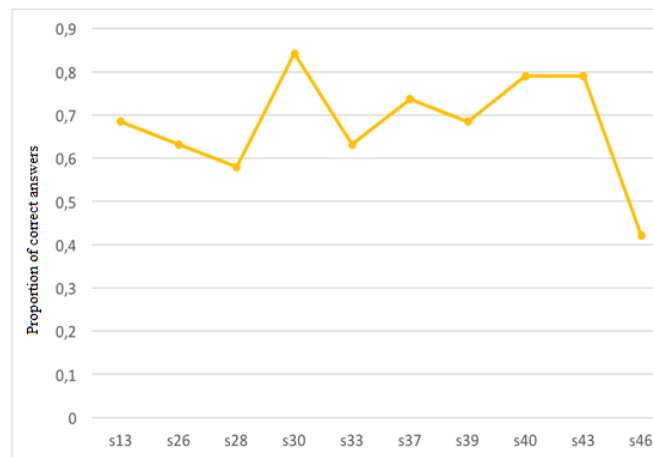
Graphic 6. Proportion of expected answers for adjunct sentences along the experiment
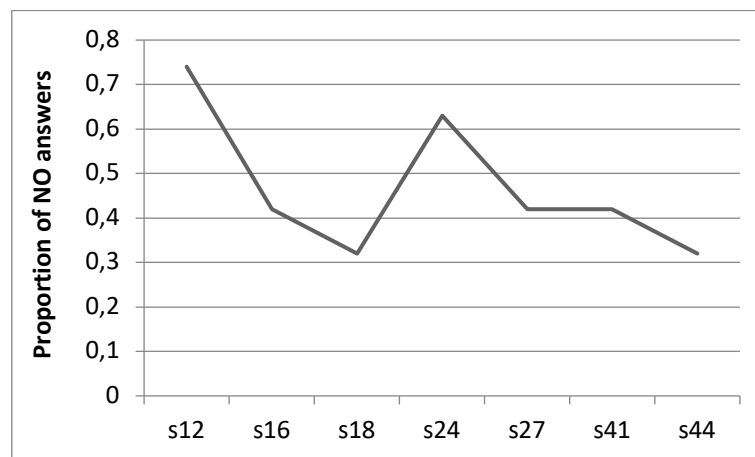
Graphic 7. Proportion of expected answers for *convencer*-sentences along the experiment

We submitted the results depicted in Graphics 5-7 as well as the repetition results shown in Graphic 4 to two tendency tests (Cox-Stuart test and Mann-Kendall test) in order to determine whether the proportion of YES- and NO-answers, as well as repetitions, showed a statistically significant tendency as the experiment unfolded. Table 2 below shows that under both tendency tests, only the decrease of repetition in the course of the experiment turned out to be statistically significant ($p = 0.006$ under Cox-Stuart test and $p = 0.001$ under Mann-Kendall test). The p-values above 0.05 for each sentence type in both tests show that there was no correlation between expected answers and order of presentation. In other words, the reliability of the answers for the different sentence types was not affected by the order in which the sentences were presented.

P-values of tendency tests for expected answers and for repetition				
	ADJUNCT	COMPLEMENT	*CONVENCER*	REPETITION
Cox-Stuart test	0.8551	0.3613	0.8551	0.006449
Mann-Kendall test	0.78592	0.7836	0.20187	0.0016477

Table 2. Expected answers per sentence type and repetition along the experiment: p-values of tendency tests

4.6. Pattern of YES- and NO-answers per sentence type and individual behavior
Figure 11 below shows the overall pattern of YES- and NO-answers for sentences involving complements, adjuncts, and *convencer*. There is no statistically significant difference between answers for *convencer*-sentences and complement sentences (p-value for YES-answers = 0.39723; p-value for NO-answers = 0.311338) or between *convencer*-sentences and adjunct sentences (p-value for YES-answers = 0.401755; p-value for NO-answers = 0.421641).
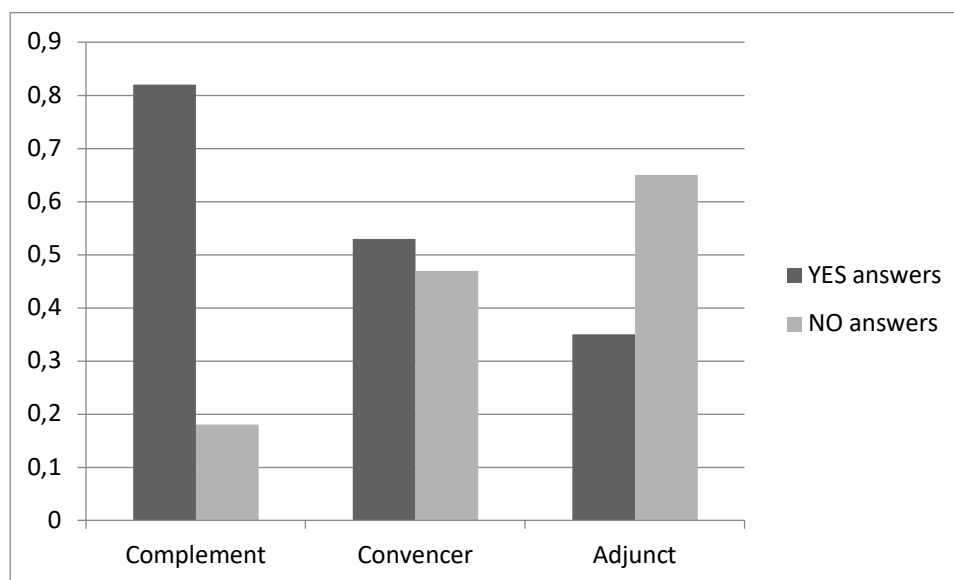

Figure 11. Proportion between YES and NO answers

Figure 11 by itself is not very illuminating. Recall that Rodrigues's (2004) took extraction of the subject of the finite clause associated with *convencer* to be quite marginal (see (8)), whereas Modesto (2011) took it to be completely acceptable (see (13a)). However, the results depicted in Figure 11 place *convencer*-sentences in a middle position between clear complements and adjuncts. In this regard, the overall performance of individual participants is not very helpful either: out of the 19 participants, 10 answered YES with respect to *convencer*-sentences most of the times and 9 answered NO.

4.7. YES- and NO-answers and syntactic complexity of the extracted wh-*phrase*
Figure 12 below depicts the proportion between YES and NO answers for the combined set of complement, adjunct and *convencer*-sentences when the type of the *wh*-phrase is taken into account. It shows that in the overall pattern, the

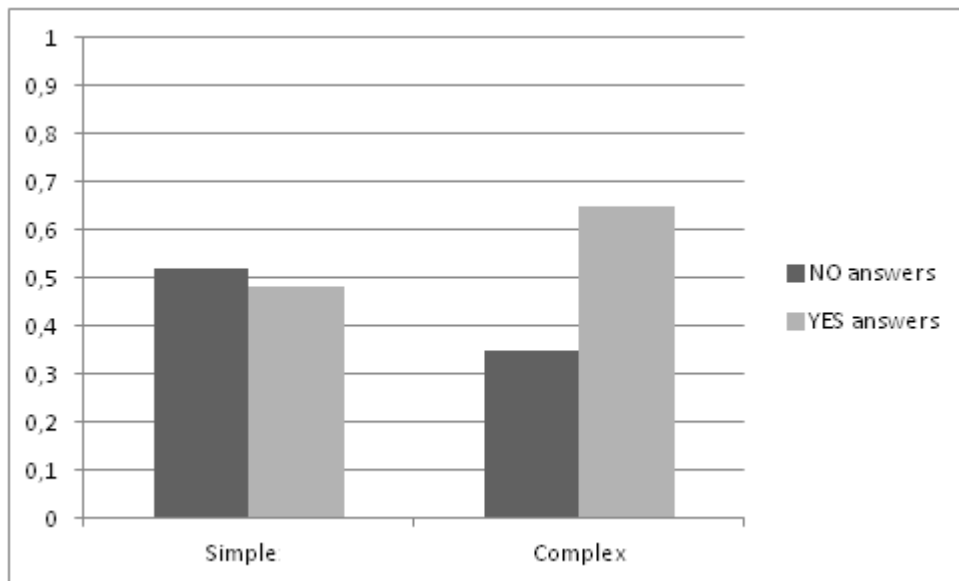correlation between complex *wh*-phrases and YES answers is statistically significant (p = 0.000238).

Figure 12. Proportion of YES and NO answers per type of *wh*-phrase

Figures 13, 14, and 15 below in turn show how each type of sentence taken in isolation interacts with the complexity of the *wh*-phrase.
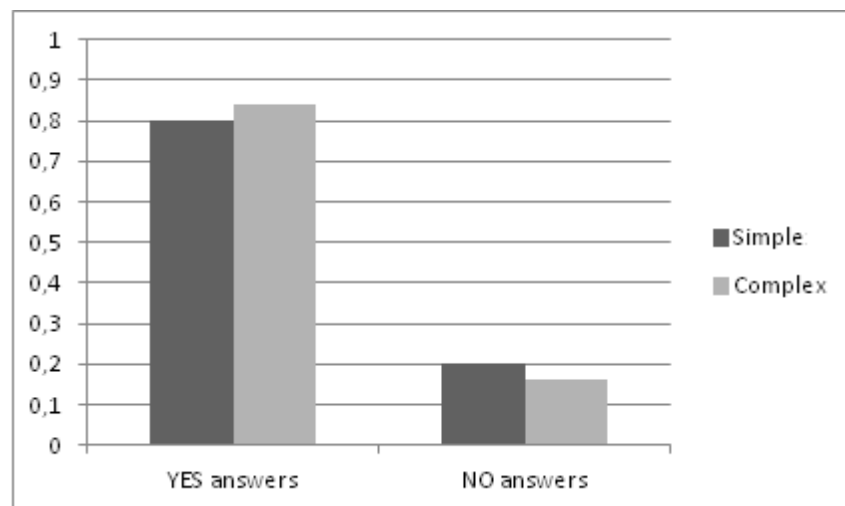

Figure 13. Proportion between YES and NO answers for complement sentences: Effect of the syntactic complexity of the *wh*-phrase
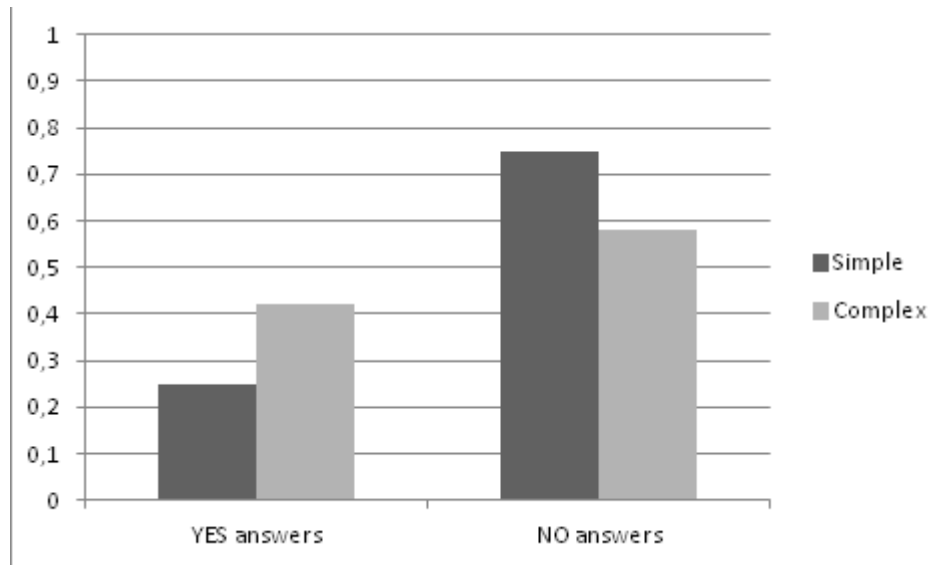
Figure 14. Proportion between YES and NO answers for adjunct sentences:
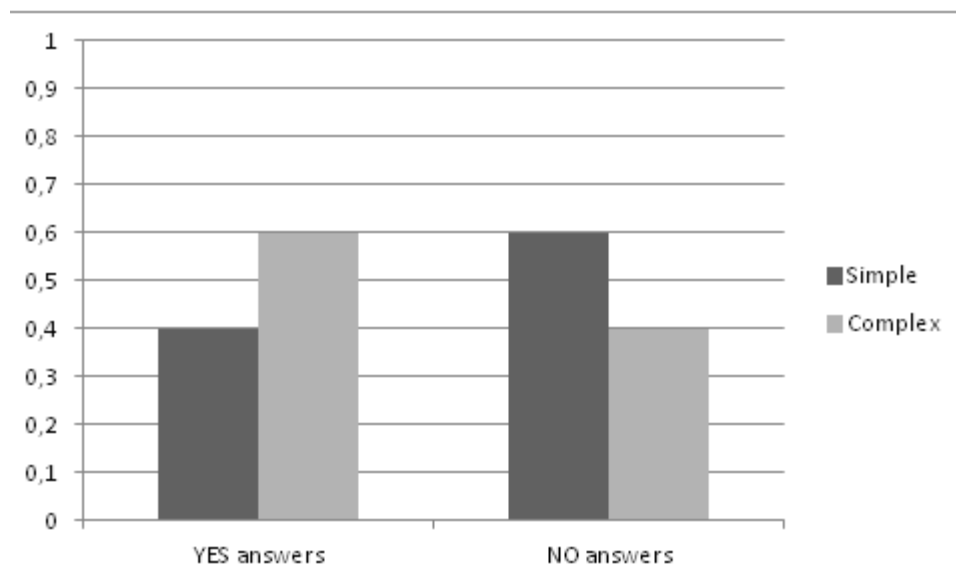Effect of the syntactic complexity of the *wh*-phrase

Figure 15. Proportion between YES and NO answers for *convencer*-sentences:
Effect of the syntactic complexity of the *wh*-phrase

Figure 13 shows that the effect of the complexity of the *wh*-phrase on extraction out of complement clauses is not statistically significant (p = 0.5706). In turn, Figure 14 shows that the complexity of the *wh*-phrase is statistically significant in the case of adjunct sentences (p = 0.01801), with complex *wh*-phrases favoring YES answers. The relevant result for our purposes is shown in Figure 15: subject extraction out of the finite clause associated with *convencer* is sensitive to the complexity of the *wh*-phrase (p = 6.801e-5), patterning like subject extraction out of adjunct clauses (see Figure 14) and not like subject extraction out of complement clauses (see Figure 13).

This result gets even more robust when we examine if there are correlations among the answers for different sentence types. We have already seen in section 4.2 that the answers for familiarization and distractor sentences did not correlate with the answers for *convencer*-sentences (see Graphics 2 and 3). In fact, the only correlation found among specific answers with respect to different sentence types involved answers for adjunct sentences and answers for *convencer*-sentences. As Graphic 8 below shows, the larger the number of YES answers a given participant offers to adjunct sentences, the larger the number of YES answers (s)he provides to *convencer*-sentences (p = 0.03619).
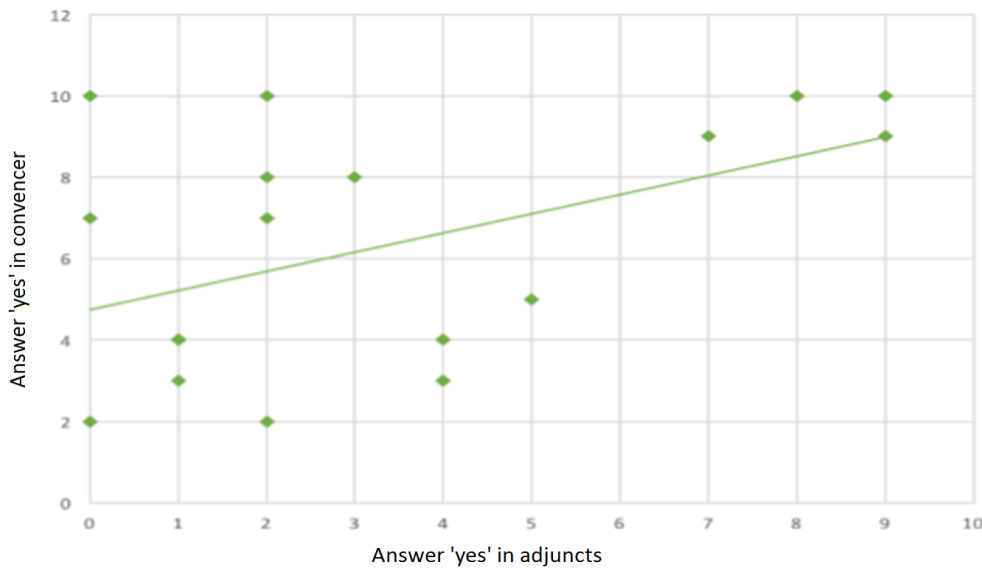
Graphic 8. Correlation between YES answers for adjunct and *convencer*-sentences

5. Discussion

At first sight, the overall pattern of YES- and NO-answers per sentence type depicted in Figure 11 and the description of the individual behavior by the participants (10 favoring YES-answers and 9 favoring NO-answers with respect to *convencer*-sentences) seem to suggest that we are simply facing a garden-variety idiolectal variation, with some speakers allowing subject extraction out of embedded clauses associated with *convencer* and some other speakers disallowing it. However, it could in principle be the case that the inconclusive picture seen in Figure 11 actually results from the interaction of each type of sentence with independent conditions.

We have seen in Figure 9 that the complexity of the *wh*-phrase was not a relevant factor for repetition. However, the descriptive content available in complex *wh*-phrases is arguably more prone to licensing a D-linked interpretation and D-linking attenuates island effects (cf. Pesetsky 1987). Thus, one could expect the complexity of the moved *wh*-phrase to have a greater impact on YES answers in adjunct sentences than in complement sentences. This expectation was indeed borne out: the complexity of the *wh*-phrase was statistically significant for extraction out of adjuncts (see Figure 14), but not for extraction out of

complements (see Figure 13). Once this contrast was established, the next step was to consider the behavior of *convencer*-sentences in this regard and the result was that the complexity of the *wh*-phrase was also statistically significant (see Figure 15). In other words, this indicates that the picture seen in Figure 11 is misleading and that *convencer*-sentences pattern like adjunct sentences and not like complement sentences.

This conclusion becomes even clearer when we take the overall contrast between Figure 13, on the one hand, and Figures 14 and 15, on the other, in the light of the results displayed in Graphic 8. Graphic 8 shows that the tolerance regarding subject extraction out of finite clauses associated with *convencer* exhibited by individual participants correlates with their tolerance regarding subject extraction out of adjunct clauses. The more tolerant a given speaker is with respect to subject extraction out of adjunct clauses, the more tolerant s/he will be with respect to *convencer*-sentences.

Another revealing result was provided by the behavior of *convencer*-sentences regarding repetition. Recall that participants were requested to say if the questions formulated by a computer software were well formed or not (see section 3). Given that they had the possibility of pressing REPETE to listen to the sentence once again before passing on their judgment, it is reasonable to think that they would take advantage of this option in cases that were harder to process or harder to judge. Furthermore, given that complements are transparent domains for extraction as opposed to adjuncts, the expectation was that adjunct sentences should be harder to process and evaluate. Therefore, the proportion of repetition was expected to be larger for adjunct sentences than for complement sentences. This expectation was indeed borne out, as seen in Figure 10. Importantly, the proportion of repetition for *convencer*-sentences (0.35/0.65) in Figure 10 is much more similar to the proportion for adjunct sentences (0.36/0.64) than complement sentences (0.26/0.74). Again, this corroborates Rodrigues's (2004) proposal that the finite clause associated with *convencer* has the behavior of an adjunct clause.

6. Conclusion

As discussed in section 2, subject extraction out of finite clauses associated with verbs like *convencer* in BP may provide a crucial basis for us to evaluate Modesto's (2000, 2011) empirical argument against analyzing definite null subjects in BP in terms of Hornstein's (1999, 2001) MTC, as advocated by Ferreira 2000, 2009, Rodrigues 2002, 2004 and Nunes 2008, 2009, among others. If the embedded clause of a sentence such as (4), for instance, repeated below in (16), turns out to function as a complement, Modesto's argument is a valid one, for movement of the embedded subject to the matrix subject position should be blocked by the intervening object. On the other hand, if it turns out to function as an adjunct, as proposed by Rodrigues (2004), Modesto's point becomes mute, for the MTC takes adjunct control to be derived via sideward movement of the subject of the will-be adjunct to the subject position of the subordinating clause (see Hornstein 1999, 2001); in this case, an object in the subordinating clause does not count as an intervener, for it does not c-command the trace of the moved subject (see (11) and (12)).

(16)   [O Paulo]$_1$ convenceu [a Maria]$_2$ que *ec$_1$/*$_2$* tinha que ir embora
           the Paulo   convinced the Maria   that   had   that go away
           'Paulo convinced Maria that he had to go away'

The results of our experiment help us understand the disagreement in the judgments regarding subject extraction in *convencer*-sentences reported in the literature (see (8) vs. (11a)). Taken in isolation, these sentences appear to fluctuate between complement and adjunct sentences as far as subject extraction is concerned and this fluctuation is also observed with respect to the participants individually, with some speakers being more tolerant than others. We have shown that this conundrum can nonetheless be unveiled if we take into account factors that may independently affect speakers' judgments regarding subject extraction.

One such factor was repetition. Under the assumption that island violations are harder to process, the sentences of the experiment that involved island violations should trigger more instances of repetition. As we see in Figure 16 (see Figure 10), *convencer*-sentences patterned like adjunct sentences with respect to repetition.
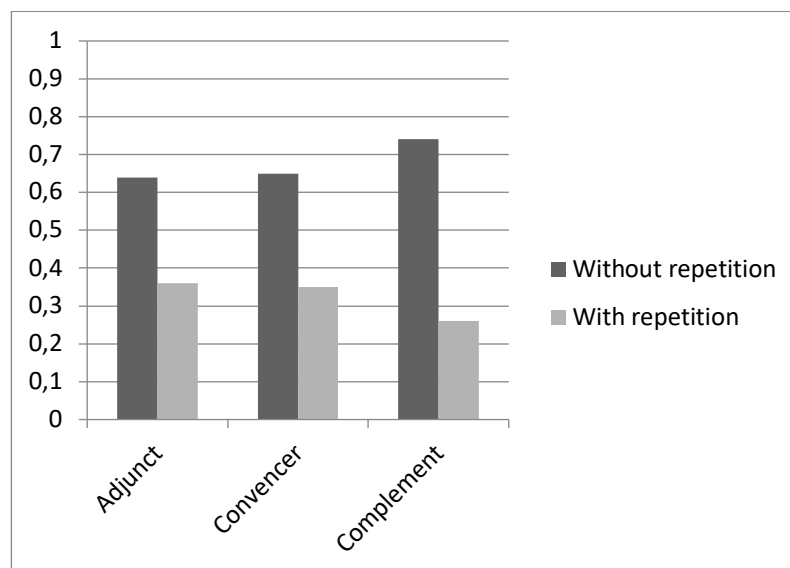

Figure 16. Proportion of repetition for complement, adjunct and *convencer*-sentences

Another factor that independently affects the acceptability of constructions involving movement is D-linking. In particular, island effects get weakened if the moved element involves a D-linked *wh*-phrase (cf. Pesetsky 1987). In the experiment, we indirectly controlled for D-linking by investigating the behavior of simple and complex *wh*-phrases for each sentence type. Figure 17 below, which portrays the results for NO-answers in the Figures 13, 14, and 15, shows that complex *wh*-phrases have a greater impact in reducing the proportion of NO answers in adjunct and *convencer*-sentences than in complement sentences. In other words, the amelioration effect that D-linking produces with respect to adjunct island violations is also observed in *convencer*-sentences.
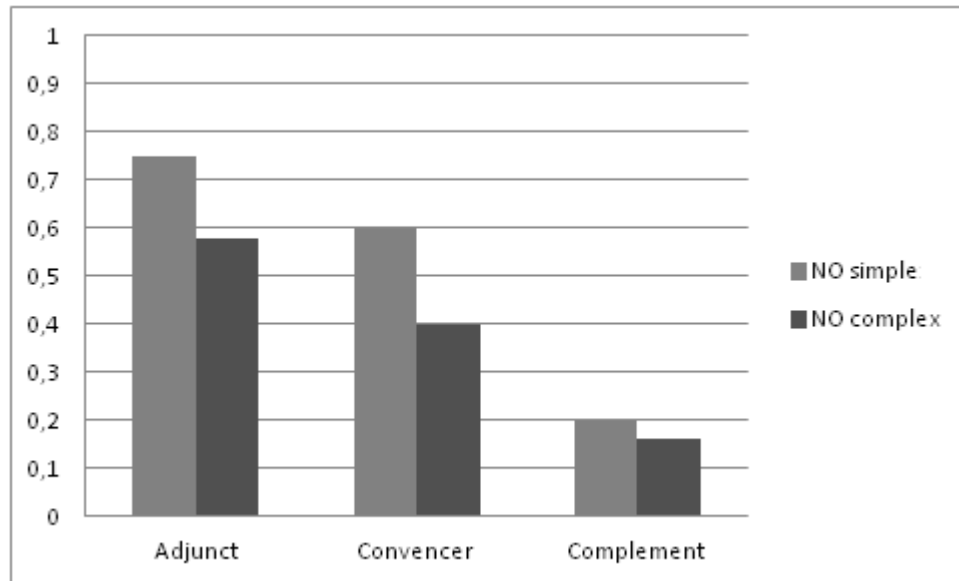
Figure 17. Effect of the syntactic complexity of the *wh*-phrases on NO answers

Finally, Graphic 8, repeated below, shows that the proportion of YES answers to adjunct sentences was correlated with the proportion of YES answers to *convencer*-sentences. This indicates that the individual variation detected among the participants regarding the acceptability of subject extraction out of *convencer*-sentences is actually a reflex of how tolerant a given individual is with respect to subject extraction out of adjunct islands.
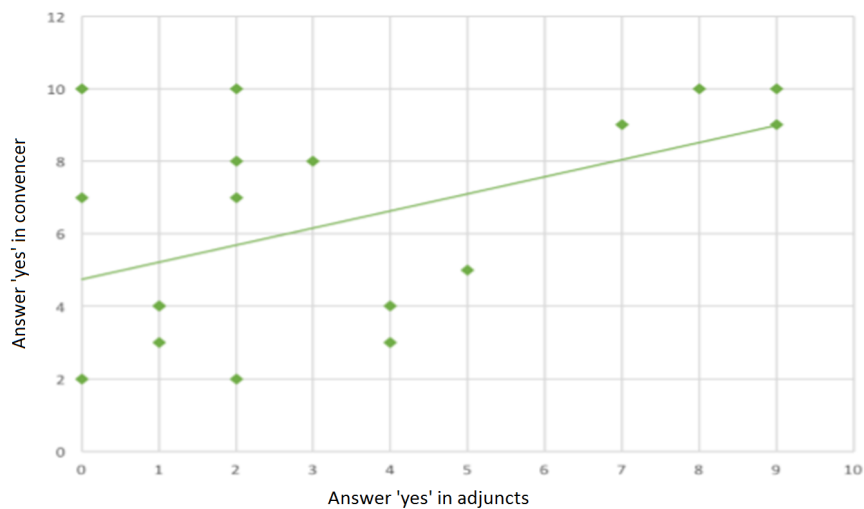
Graphic 8. Correlation between YES answers for adjunct and *convencer* sentences

These results support Rodrigues's (2004) proposal that the finite embedded clause associated with *convencer* in BP is not a real complement but some sort of adjunct, thus invalidating Modesto's (2000, 2011) argument. The conclusion is that subject orientation in sentences like (16) does not provide empirical evidence against the movement approach to definite null subjects in BP,

for it can be captured in terms of sideward movement, as argued for by
Rodrigues's (2004).

References

Barbosa, Pilar, M. Eugênia Duarte and Mary A. Kato. 2005. "Null subjects in
European and Brazilian Portuguese". Journal of Portuguese Linguistics 4:
11-52.
http://dx.doi.org/10.5334/jpl.158

Chao, Wynn. 1983. "The interpretation of null subjects: Brazilian Portuguese".
Cahiers Linguistiques d'Ottawa 11: 69-74.

Chomsky, Noam. 1995. The minimalist program. Cambridge, MA: MIT Press.

Coelho, Claudia. In progress. Controvérsias sobre o Sujeito Nulo do Português
Brasileiro: Um Estudo Experimental. M.A. thesis, Universidade de São
Paulo.

Duarte, Maria Eugênia. 1995. A Perda do Princípio "Evite Pronome" no
Português Brasileiro. Ph.D. thesis, Universidade Estadual de Campinas.

Ferreira, Marcelo. 2000. Argumentos Nulos em Português Brasileiro. M.A. thesis,
Universidade Estadual de Campinas.

Ferreira, Marcelo. 2009. "Null subjects and finite control in Brazilian
Portuguese". In J. Nunes (ed.), Minimalist essays on Brazilian Portuguese
syntax, 17-49. Amsterdam/Philadelphia: John Benjamins.
http://dx.doi.org/10.1075/la.142.04fer

Figueiredo Silva, Maria Cristina. 1996. A posição do sujeito no português
brasileiro: frases finitas e infinitivas. Campinas: Editora da UNICAMP.
http://dx.doi.org/10.18309/anp.v2i5.313

Galves, Charlotte. 1987. "A sintaxe do português brasileiro". Ensaios de
Linguística 13: 31-50.

Galves, Charlotte. 2001. Ensaios sobre as gramáticas do português. Campinas:
Editora da UNICAMP.

Holmberg, Anders, Aarti Nayudu, & Michelle Sheehan. 2009. "Three partial null-
subject languages: A comparison of Brazilian Portuguese, Finnish and
Marathi". Studia Linguistica 63: 59-97.
http://dx.doi.org/10.1111/j.1467-9582.2008.01154.x

Hornstein, Norbert. 1999. "Movement and control". Linguistic Inquiry 30: 69-96.

Hornstein, Norbert. 2001. Move! A minimalist theory of construal. Oxford:
Blackwell.

Kato, Mary A. 1999. "Strong pronouns, weak pronominals and the null subject
parameter". Probus 11: 1–37.
http://dx.doi.org/10.1515/prbs.1999.11.1.1

Kato, Mary A. & Esmeralda Negrão (eds.). 2000. Brazilian Portuguese and the
null subject parameter. Madrid & Frankfurt am Main: Iberoamericana &
Vervuert.

Modesto, Marcelo. 2000. On the Identification of Null Arguments. Ph.D. thesis,
University of South California.

Modesto, Marcelo. 2011. "Finite control: Where movement goes wrong in
Brazilian Portuguese". Journal of Portuguese Linguistics 10: 3-30.
http://dx.doi.org/10.5334/jpl.95

Moreira da Silva, Samuel. 1983. Études sur la Symétrie et L'asymétrie Sujet/Objet dans le Portugais du Brésil. Ph.D. thesis, Université de Paris VIII.

Negrão, Esmeralda. 1986. Anaphora in Brazilian Portuguese Complement Structures. Ph.D. thesis, University of Wisconsin.

Nunes, Jairo. 2001. "Sideward movement". Linguistic Inquiry 31: 303-344. http://dx.doi.org/10.1162/00243890152001780

Nunes, Jairo. 2004. Linearization of chains and sideward movement. Cambridge, MA: MIT Press.

Nunes, Jairo. 2008. "Inherent Case as a licensing condition for A-movement: The case of hyper-raising constructions in Brazilian Portuguese". Journal of Portuguese Linguistics 7: 83-108. http://dx.doi.org/10.5334/jpl.129

Nunes, Jairo. 2009. "Dummy prepositions and the licensing of null subjects in Brazilian Portuguese". In E. Aboh, E. van der Linden, J. Quer & P. Sleeman (eds.), Romance languages and linguistic theory, 243-265. Amsterdam/Philadelphia: John Benjamins. http://dx.doi.org/10.1075/rllt.1.13nun

Nunes, Jairo. 2011. "On the diachronic reanalysis of null subjects and null objects in Brazilian Portuguese". In E. Rinke and T. Kupish (eds.), The development of grammar: Language acquisition and diachronic change - In Honor of Jürgen M. Meisel, 331-354. Amsterdam/Philadelphia: John Benjamins. http://dx.doi.org/10.1075/hsm.11.18nun

Nunes, Jairo. 2013. "Edge features legitimando movimento-A". ReVEL, edição especial 7: 35-50.

Nunes, Jairo. 2014. "Adjunct control and edge features". In P. Kosta, S. Franks, T. Radeva-Bork & L. Schürcks (eds.), Minimalism and beyond: Radicalizing the interfaces, 79-108. Amsterdam/Philadelphia: John Benjamins.

Nunes, Jairo. 2016. "Edge features on phase heads or moving elements?". Florida Linguistics Papers 3: 1-18.

Pesetsky David.1987. "Wh-in-Situ: movement and unselective binding". In E. Reuland & A. Meulen (eds.), The representation of (in)definiteness, 98-129. Cambridge, MA: MIT Press.

Petersen, Maria Carolina. 2011. O Licenciamento do Sujeito Nulo em Orações Subjuntivas no Português Brasileiro: Contribuições para a Teoria de Controle por Movimento. M.A. thesis, Universidade de São Paulo.

Rodrigues, Cilene. 2002. "Morphology and null subjects in Brazilian Portuguese". In D. Lightfoot (ed.), Syntactic effects of morphological change, 160–178. Oxford & New York: Oxford University Press.

Rodrigues, Cilene. 2004. Impoverished Morphology and A-movement out of Case Domains. Ph.D. thesis, University of Maryland. http://dx.doi.org/10.1093/acprof:oso/9780199250691.003.0009

Saab, Andrés, 2016. "On the notion of partial (non-) pro-drop in Romance". In M. A. Kato & F. Ordoñez (eds.): The morphosyntax of Spanish and Portuguese in Latin America, 49-77. Oxford: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780190465889.003.0003