https://dx.doi.org/10.1590/1678-460X202339251209

Articles

A Survey on Generic Impersonal Structures in Brazilian Portuguese[1]

Uma pesquisa sobre estruturas impessoais genéricas do português brasileiro

Karina Gomes Bertolino[2]

ABSTRACT

This article presents a survey on null generic impersonal structures conducted with native speaker of Brazilian Portuguese (BP). The objective of the survey was (i) to investigate whether null impersonal constructions with the generic reading need to have elements such as fronted adverbials to check the EPP, a hypothesis put forward by Holmberg (2005, 2010a,b); (ii) verify to what extent BP-speakers require an overt marker of genericity (e.g., the generic clitic *se* and overt necessity deontic modals) in sentences with null generic pronouns, as suggested by Bertolino (2017). The survey consisted of an Acceptability Judgment Test, conducted online. All sentences were presented with a context favoring the generic reading of the null subject. The results of this survey show that (i) the EPP does not need to be satisfied in null impersonal constructions in BP; (ii) although

1. I want to express my gratitude to Diane Lillo-Martin, William Snyder and Željko Bošković for their careful reading and corrections. Thanks also to two anonymous reviwers for D.E.L.TA, who provided helpful comments and sugestions. All remaining errors are my own.
2. Departamento de Linguística, Universidade de São Paulo. São Paulo – Brasil. https://orcid.org/0000-0001-5089-4688. E-mail: karina.bertolino@usp.br.

BP-speakers prefer null generic impersonal constructions with overt markers of genericity, they do not require it.

Keywords: *null impersonals; genericity; EPP.*

RESUMO

Este artigo apresenta uma pesquisa conduzida com falantes nativos de português brasileiro (PB) sobre estruturas impessoais nulas com leitura genérica. O objetivo da pesquisa era (i) investigar se elementos como advérbios fronteados são necessários para satisfazer o EPP nessas estruturas, uma hipótese que é aventada por Holmberg (2005; 2010a,b); (ii) verificar até que ponto falantes de PB exigem uma marca expressa de genericidade (como o clítico genérico *se* e modais deônticos de necessidade) em sentenças com pronomes nulos genéricos, conforme sugerido por Bertolino (2017). A pesquisa, conduzida on-line, utilizou a Tarefa de Julgamento de Aceitabilidade como metodologia. Todas as sentenças foram apresentadas com um contexto favorecendo a leitura genérica do sujeito nulo. Os resultados desta pesquisa mostram que (i) o EPP não precisa ser satisfeito em estruturas impessoais nulas do PB; (ii) embora falantes nativos de PB prefiram marcas expressas de genericidade em estruturas impessoais com sujeito nulo genérico, eles não as requerem.

Palavras-chave: *impessoais nulas; genericidade; EPP.*

1. Introduction

In this paper, I report the results of an online survey conducted with native speakers of Brazilian Portuguese (BP) on null generic impersonal structures, that is, sentences in which the null subject means "people in general", as in the sentence in (1). This is a descriptive study which does not aim to provide an in-depth analysis of the collected data.

Respondents rated auditorily-presented sentences given in context on a 5-point scale (to be described in more detail below). The main purpose of the survey was to see if the hypotheses in (i) and (ii) were supported.

(1) a. Ø Não pode largar a escola.
not can:3SG drop:INF the school
'One should not drop out of school.'
b. Nessa escola Ø não pode comer banana.
in.this school not can eat:INF banana
'In this school one cannot eat bananas.'

Hypothesis

(i) In null generic impersonal constructions, the EPP needs to be checked by an element other than the null pronoun which values T's φ-features.

(ii) The generic reading of null impersonal sentences requires an overt marker of genericity.

The hypothesis in (i) is raised by Holmberg (2005, 2010a, b). According to the author, as null generic pronouns are φP pronouns which are the tail of a chain, they are unable to check the EPP. As such, the EPP should be checked by an element other than the null subject. This hypothesis predicts that if there is in fact an EPP feature that needs to be checked, speakers should not accept sentences such as (1a), which have the generic reading but no element to check the EPP.

Regarding the hypothesis in (ii), as generic operators are modals of necessity (Heim, 1982), it is plausible to formulate the hypothesis that this modal has to be overtly represented in generic impersonal structures, at least in some languages. Applied to BP, this hypothesis predicts that the sentence (1b) should be acceptable, since it has an overt deontic necessity modal acting as a genericity marker.

The hypothesis in (ii) was first tested by Bertolino (2017) in a pilot study conducted with native speakers of BP. The results of this pilot study suggested that native speakers of BP only accept generic null impersonal structures with an overt marker of genericity (such as the generic *se* and a necessity deontic modal).[3][4][5]

3. The generic reading of *se* is associated with a non-specific/non-bounded time reference. This same clitic can have the existential reading when associated with a specific/bounded time reference, such as the past tense and perfective aspect (Cinque, 1988; D'Alessandro, 2007). Therefore, it is not only the presence of *se* which makes a sentence generic, but the combination of *se* and a non-specific/non-bounded time reference.

4. The acceptability of null impersonals with necessity epistemic modals was not tested by Bertolino (2017) and in the survey to be reported here. Given the empirical coverage of these two studies, I restrict myself in claiming throughout this article that necessity deontic modals act as overt markers of genericity, instead of claiming that all necessity modals act as an overt marker of genericity.

5. As one of the reviewers pointed out, there are other strategies to express indetermination in Brazilian Portuguese besides the use of *se* and null subjects (e.g., the use of the pronoun *você* 'you'). See Vargas (2010, 2012), Othero & Miranda da Silva (2020) for details.

In this study, in order to assess the hypothesis in (ii), generic structures with a deontic modal or generic *se* were compared with generic structures without these elements. Under (ii), the prediction was that generic structures without an overt marker of genericity would not be acceptable.

This article is organized as follows: in section 2, I explain the methodology and the materials adopted in the survey. Section 3 describes the participants in the study. In section 4, I report the results of the survey (descriptive statistics). Section 5 separately discusses the results of some of the sentences in the survey which had the existential reading instead of the generic one. Section 6 discusses whether our data support the hypotheses in (i) and (ii) (inferential statistics). Section 7 concludes the article.

2. Methodology and Material

The survey consisted of an Acceptability Judgment Test, conducted online. The survey was hosted by *Qualtrics*, a web-based platform that allows researchers to create surveys and generate reports. *Qualtrics* supports many languages, including Portuguese, in such a way that I was able to run this survey using only the participants' native language.

Table 1 shows the test sentences that were used in the survey together with their code representing the type of sentence (code) (column 1) and the code explanation (column 2) (e.g., ADV+D (code), locative adverbial PP + deontic modal (code explanation)). Additionally, five fillers were presented randomly in the survey. One of the fillers was categorically ungrammatical, two of them were completely grammatical and two of them admitted gradation in the judgement. In order to keep the survey short and increase the rate of completion, no more than 19 sentences were used. The average time to complete the survey was 13 minutes. Notice that there were two sentences (in the conditions ADV and NoADV) in the survey that were compatible with the existential reading of the null subject: they will be analyzed separately (see section 5). See appendix for all the sentences that were used in the survey.

Table 1 – sentences used in the survey[6]

Code	Code Explanation	Sentences
ADV	Locative adverbial PP	Nessa loja (se) vende saia. In.this store sell:3SG skirt *Null subject:* 'In this store someone sells skirts.' *Se:* 'In this store one sells skirt.'
		Nessa universidade (se) estuda na biblioteca. In.this university study:3SG in.the library 'In this university one studies in the library.'
NEG	Negation	Não (se) faz bolo de cenoura desse jeito. Not make:3SG cake of carrot in.this way 'One doesn't make carrot cake in this way.'
		Não (se) vende mais geladinho. Not sell:3SG more ice.pop 'One doesn't sell ice pops anymore.'
ADV+NEG	Locative adverbial PP + Negation	No Brasil não (se) colhe cogumelo em bosques. In.the Brazil not pick:3SG mushroom in forests 'In Brazil one does not pick mushrooms in forests.'
		Nos Estados Unidos não (se) assiste novela. In.the States United not watch:3SG soap.opera 'In the US one does not watch soap operas.'
ADV+D	Locative adverbial PP + Deontic Modal	Nessa escola (se) tem que comer verdura na hora do almoço. In.this school have:3SG that eat vegetables at.the time of lunch 'In this school one has to eat vegetables at lunch time.'
		Nessa escola (se) tem que entregar lição de casa. In.this school have:3SG that turn.in:INF lesson of home 'In this school one has to turn in homework assignments.'
NEG+D	Negation + Deontic Modal	Não (se) pode largar a escola. Not can:3SG leave:INF the school 'One cannot drop out of school.'
		Não (se) pode andar de bicicleta na calçada. Not can walk:INF of bicycle in.the sidewalk 'One cannot ride a bicycle on the sidewalk.'
ADV+NEG+D	Locative adverbial PP + Negation + Deontic Modal	Nessa escola não (se) pode trazer animal de estimação. In.this school not can:3SG bring:INF pet 'In this school one cannot bring pets.'
		Em Samoa não (se) pode esquecer o aniversário da esposa. In Samoa not can:3SG forget:INF the birthday of.the wife 'In Samoa one cannot forget one's wife's birthday.'
NoADV	No Adverbial	(Se) vende doce. sell:3SG dessert *Null subject:* 'Someone sells desserts.' *Se:* 'In this store one sells desserts.'
		(Se) toma banho todo dia. take:3SG shower every day 'One shower every day.'

6. Where I indicated that the existential reading of the pronoun is possible (by using the pronoun *someone* in the translation), it only happens in null impersonals, not in impersonals with *se*, as I assume that the clitic *se* can only have a generic reading when associated with a non-specific/non-bounded time reference (see footnote 3).

It is important to mention that all sentences were presented with a context initially given, as the example below shows:

(2) The world is full of husbands with poor memory, but Samoa, a country in Oceania, is trying to change it! In Samoa a husband is forbidden to forget his wife's birthday. If he does, he is brought to justice and has to indemnify his wife for his actions.

(3) Em Samoa não (se) pode esquecer o aniversário da esposa.
In Samoa not can:3SG forget:INF the birthday of.the wife
'In Samoa one cannot forget one's wife's birthday.'

The participants listened to an audio with the context and the sentence in BP (no transcription was provided). The participants were informed in the initial screen that they had to judge the form of the sentence and not its content, providing judgements on a scale from 1 (very bad) to 5 (perfect). Half of the participants listened to test sentences with *se* and the other half listened to identical test sentences, but without *se.*

For each one of the structures presented in Table 1, I summarize below which judgements were predicted under the hypotheses (i) and (ii) for the sentences *with se* and *without se*:

ADV: under the hypothesis in (i), both the sentences *with se* and *without se* should be judged as grammatical, since the adverbial can check the EPP. However, under the hypothesis in (ii), we expected that sentences *with se* would be judged as better than sentences *without se*, since I predict that the generic reading of impersonal sentences in BP requires an overt marker of genericity.

NEG: sentences in this condition should be judged as ungrammatical, since there is no element to check the EPP (hypothesis (i)).

ADV+NEG: under the hypothesis in (i), sentences in this condition should be judged as grammatical, since the adverbial can check the EPP. Under the hypothesis in (ii), sentences *with se* would be judged as better than sentences *without se* for exactly the same reason as in ADV: BP-speaking adults prefer generic structures with an overt marker

of genericity than without it. As *se* is an overt marker of genericity, ADV+NEG is expected to be better with *se* than without it.

ADV+NEG+D: under the hypothesis in (i), sentences in this condition should be judged as grammatical, as the fronted adverbial can check the EPP. Under the hypothesis in (ii), sentences *without se* in this condition should be judged as equally good as sentences *with se* if nothing in the language prevents double markers of genericity (*se* and the modal).

ADV+D: under the hypothesis in (i), sentences in this condition should be judged as grammatical, since the adverbial can check the EPP. Under the hypothesis in (ii), sentences *without se* in this condition should be judged as equally good as sentences *with se* for the same reason as in ADV+NEG+D.

NEG+D: sentences in this condition should be judged as ungrammatical, since there is no element to check the EPP (hypothesis (i)).

NoADV: sentences in this condition should be judged as ungrammatical, since there is no element to check the EPP (hypothesis (i)).

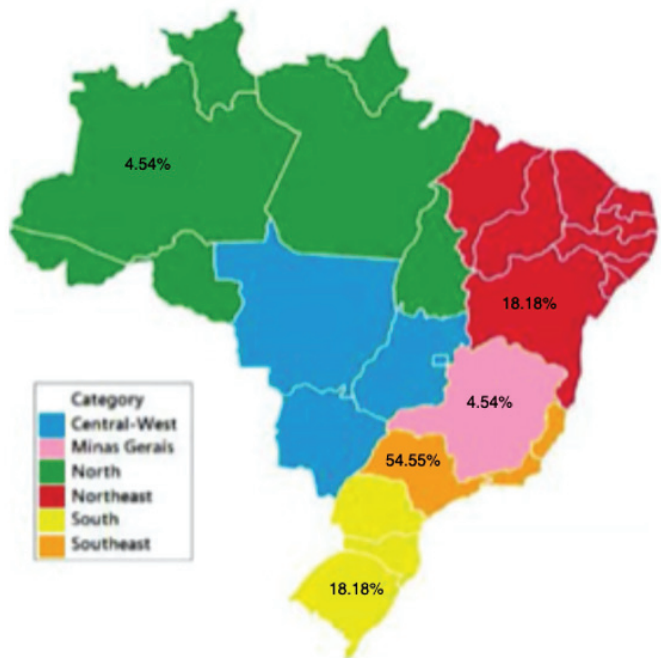
3. Participants

A total of 22 participants had their answers analyzed. I excluded 3 participants that demonstrated lack of seriousness in the way they filled in the survey (two participants were excluded for judging all sentences, including fillers, as either very bad or perfect, and one participant was excluded because they took about 1 minute to complete the survey, which shows that they did not listen to the audios before providing a judgment). 11 of the participants included in the analysis listened to sentences *with se* (Test 1) and the other 11 participants listened to the sentences *without se* (Test 2). The participants included in the analysis responded to the fillers as expected.

The participants were from many regions of Brazil, although the distribution is clearly unbalanced, with a predominance of participants

from the Southeast region of Brazil (both for Test 1 and 2), as can be seen in the map below.[7] [8]As the number of participants per region is small and unbalanced, I will not report the results by region.

Graph 1 – Participant's geographic distribution



Participants were between the ages of 19 to 46. The mean age was 31.73 for those who took Test 1 (sentences *without se*) and 29.82 for those that took Test 2 (sentences *with se*). There is no significant difference in the age of participants that received Test 1 (*without se*)

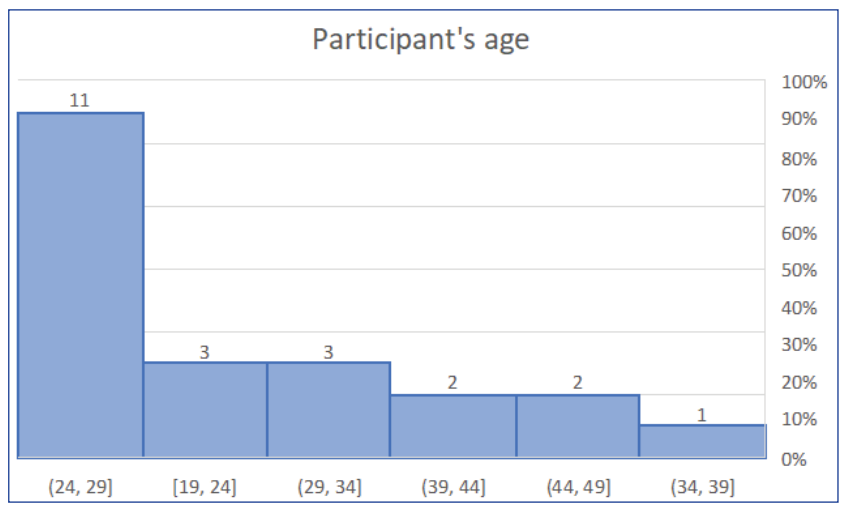
7. The percentage refers to the proportion of participants in each region, not in particular states. For example, even if the number 54.55% is in the state of São Paulo on the map, we are counting the proportion of participants in the whole Southeast region represented in orange.

8. Minas Gerais is a state in the Southeast region, represented in pink in our map. The initial idea was to represent Minas Gerais separately, due to the particularity of their dialect which does not align with the rest of the Southeast region. This dialect is known to be innovative as regards null *vs* overt subjects in particular, and its loss of *se*. However, only one participant out of 22 was from Minas Gerais. Therefore, it is impossible to extract any generalization about this dialect.

DELTA
39.2
2023

and Test 2 (*with se*): U = 57.5, p = 0.843 (2-tailed) (Mann-Whitney test). The histogram below shows the distribution of the participant's ages. As we can see, most of the participants were between 24- and 29-years-old (11 out of 22).[9]

Graph 2 – Histogram of age of participants

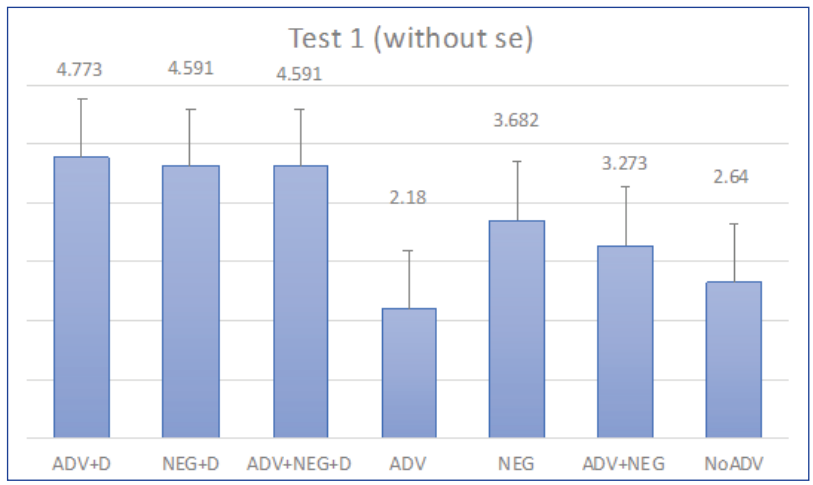


4. Descriptive Statistics

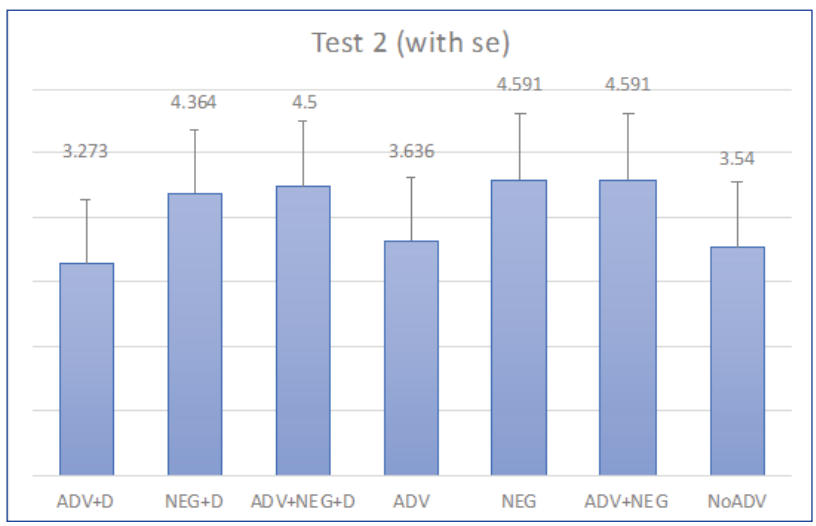
The bar-graphs below show the structures tested and their respective means on a 1 (very bad) to 5 (perfect) scale. The bar-graph excludes the sentences in the condition ADV and NoADV which had a context favoring the existential reading. These sentences are discussed separately in section 5.

9. As one of the reviewer pointed out, generality speaking, the young Brazilian population tend to accept and use constructions without *se*. Consequentely, age could be a factor affecting the results. However, as only a few of our participants were of older age groups (e.g., in their forties), it is hard to evaluate whether the age of the participants had any effect on the results.

Graph 3 – Average for each structure: sentences without se (Test 1)



Graph 4 – Average for each structure: sentences with se (Test 2)



The following table shows the means and standard deviations for the filler items. As expected, the sentence which was categorically ungrammatical received a very low score (≈ 1), while the two grammatical sentences received high scores (≈ 4-5). The ones that admitted gradation were judged as in between (≈ 3-4). The fillers were exactly the same for both Test 1 (*without se*) and Test 2 (*with se*)

Table 2 – Averages and standard deviation for filler items

Grammaticality	**Test 1 (without *se*)**		**Test 2 (with *se*)**	
Average (AV) / (SD)	AV	SD	AV	SD
?	4.09	0.67	3.55	1.30
?	4.18	1.03	3.09	1.38
✓	5.00	0	4.45	0.89
✓	4.64	0.88	4.91	0.29
✗	1.18	0.57	1.45	1.16

I will first discuss the results for Test 1 (*without se*). As seen in Graph 3, sentences with an overt deontic necessity modal and a null subject were judged as better than sentences without the modal. All the stories or contexts that preceded the test sentences had a lawlike background or described a tradition, or common behavior of people in a certain society (e.g., In Brazil people don't pick mushroom in forests). These were sentences that had necessarily a generic reading, as can be seen in the example in (4):

(4) a. Brazilian soap operas are famous around the world, but this popularity did not reach every country. In 1995, the Yugoslav War stopped for 1 week so that the country could watch the last episodes of 'Slave Isaura'! But, as said before, it is not everywhere that people watch soap operas.

b. Nos Estados Unidos Ø não assiste novela.
in.the States United not watch:3SG soap.opera
'In the United States **one** doesn't watch soap operas.'

It would be reasonable to think that speakers gave a lower value in the 1 to 5 scale to the sentences in the conditions ADV, ADV+NEG and NEG compared to the sentences in the conditions ADV+D, ADV+NEG+D and NEG+D because generic sentences with an overt marker of genericity are at least preferred by speakers, but notice that they are not completely unacceptable. As ADV, ADV+NEG and NEG sentences lack an overt modal, the generic reading, in general, was not considered "perfect" by the participants. In this way, we could explain the difference seen in the results: sentences with a deontic modal received a high value in the judgement scale because they had a modal overtly marking the genericity of the sentences; on the other hand, sentences without a deontic modal received a lower value in

the judgment scale because speakers needed an overt expression of genericity in these sentences.

The fact that the average score for the condition NEG+D was high (4.591) suggests that the EPP does not hold, at least for BP. There is no element in sentences in this condition that could check the EPP. Notice that NEG+D received the same score as ADV+NEG+D (4.591), a sentence in which the EPP could be checked by the adverbial, as suggested in Holmberg (2005; 2010a,b), Rodrigues (2004) and Holmberg et al. (2009).

Regarding the results for Test 2 (*with se*) (Graph 4), it is important to remember that the context that preceded these sentences was the same as the one that preceded sentences *without se*. The test sentences were also identical, the only difference being that these sentences had *se*, as we can see by the contrast between sentences (5) (Test 1) and (6) (Test 2):

(5) Ø Não vende mais geladinho.
not sell:3SG more freeze.pop
'They don't sell freeze pops anymore.'

(6) Não **se** vende mais geladinho.
not SE sell:3SG more freeze.pop
'They don't sell freeze pops anymore.'

Considering the hypothesis in (i), we were expecting that the condition NEG (with *se* as an overt marker of genericity) in Test 2 would receive a low average score, in the same way as NEG+D (with the modal as an overt marker of genericity) in Test 1. In fact, the average score for the condition NEG (with *se*) was exactly the same for NEG+D (without *se*), as expected if the hypothesis that the EPP holds is not supported.

As is shown in Graph 4, the condition ADV+D in Test 2 received a lower average score compared to the other conditions with a modal: 3.27. We could not explain the low score obtained for the condition ADV+D by saying that the modal cannot co-occur with *se*, because otherwise we would expect a low score for the conditions NEG+D and ADV+NEG+D, which did not happen: these two conditions received

the near perfect scores of 4.36 and 4.5, respectively. Therefore, *se* can co-occur with the modal. Why did the condition ADV+D receive a low score then?

The condition ADV+D was the only one in which the modal *ter que* ('have to') was used. All the other conditions with a modal (NEG+D and ADV+NEG+D) had the negated form of the modal *pode* ('can'): *não pode* ('cannot'). It may be the case that there is some property in *tem que* ('have to') that disfavors its co-occurrence with *se*.

The modal *tem que* ('have to') is not semantically equivalent to *não pode* ('cannot'), the other modal used in the survey, although both of them can be understood as necessity modals. In the example in (7), *tem que* describes that in the school one has the *obligation* to bring pets. In the example in (8), *não pode* describes that in the school one is *not allowed* to bring pets:

(7) Nessa escola *e* tem que trazer animais de estimação.
in.this school *e* have:3SG that bring:INF pets
'In this school one has to bring pets.'

(8) Nessa escola *e* não pode trazer animais de estimação.
in.this school not can:3SG bring:INF pets
'In this school one cannot bring pets.'

While I am pointing out a contrast between the two modals that could explain the difference found in the results of the survey (Test 2), the precise reason why a modal such as *tem que* cannot co-occur with *se* needs to be further investigated. Also, it would be necessary to have more solid evidence that *se* is in fact disfavored with *tem que.* As an online search for academic texts shows, *tem que* can co-occur with *se* at least in the formal written register:

(9) Os defensores do JIT apontam que estoques altos se faz necessário
the defenders of.the JIT point.out that stocks high SE make:3SG necessary
quando **se tem que** esconder problemas maiores na empresa.
when SE have that hide:INF problems bigger in.the company
'The defenders of the JIT point out that high stock prices are necessary when one has to hide bigger problems in the company.'

Ferreira (2017: 26)

(10) A formação de iniciantes na arqueologia nesse ambiente pode ser, portanto,
The training of beginners in.the archeology in.this environment might be therefore
bastante inadequada, pois o que **se tem que** aprender é a produzir ciência […]
quite inadequate because the what SE have that learn:INF is to produce science
'Therefore, in this environment, the training for beginners in archeology might be quite inadequate because what one has to learn is to produce science […].'
Funari (2000: 82)

Another possible reason why participants were reluctant to accept the co-occurrence of *se* and *tem que* might be because *se* appeared as a proclisis in the test-sentences with *tem que* (i.e., *se tem que*), while the most natural structure for the speakers might be the one in which *se* appears as an enclisis (i.e., *tem que se*). This is in accordance with the judgement provided by one of the reviewers.

In section 6, I analyze these results more closely, in the attempt to investigate whether the data support the hypotheses formulated in the introduction of this article. But, before that, I will discuss the results for the sentences with the existential reading.

5. Impersonals with Existential Reading

The existential reading of the null subject pronouns in impersonal sentences in BP is possible when the sentence does not express a generalization or a general law (Carvalho 2018, 2019; Bertolino, 2020). One of the sentences in the condition ADV and another sentence in the condition NoADV admitted the existential reading or even the definite one if the reference of the pronoun would be judged as "Joaquim's mother" (11) and "Mariana's father" (12): this could be possible if the owners of the store are the ones who sell the products. This reading, though, should be considered degraded, as in BP the definite reading of the third person pronoun in matrix clauses is usually only possible in subjectless replies or when the reference of the null subject can be recovered by the physical presence of the referent (Bertolino 2020). The sentences could be analyzed as existential if some unspecified person sells the products in the stores. In this case, ADV should be judged as acceptable and NoADV not, as existential sentences require an obligatory event argument (Carvalho 2018).

DELTA
39.2
2023

(11) a. Joaquim's mother has a store called 'Just skirt.'
b. Nessa loja vende saia.
In.this store sell:3SG skirt
'In this store someone sells skirts.'

(12) a. Mariana's father has a store called 'Just dessert.'
b. Vende doce.
Sell:3SG dessert
'Someone sells desserts.'

The sentence in (11) (ADV) (Test 1, *without se*) received the average score of 4.91 (SD = 0.29). Assuming that the definite reading is not possible in (11), the existential reading was well-accepted. The sentence in (12) (NoADV) (Test 1, *without se*) received the average score of 3.64 (SD = 1.30). The results of the Wilcoxon signed-rank test show that the difference is significant ($Z = -2.8031$, $p < 0.05$, 2-tailed). This result is predicted under the assumption that impersonals with existential null subjects require an event argument.

Regarding the results for the same sentences accompanied by a context which favored the existential reading, ADV (13) and NoADV (14), in Test 2 (*with se*), they received the average score of 4.09 (SD = 0.94) and 3.09 (SD = 1.14), respectively. The results of the Wilcoxon signed-rank test show that the difference is not statistically significant ($Z = -1.5401$, $p > 0.05$, 2-tailed). Recall that it was pointed out that the impersonal clitic *se* is associated with the generic reading when the sentence is in the present tense and imperfective aspect (see footnote 3). As the context favoring the existential reading is incongruent with the generic reading of the sentences by themselves, one would expect that both sentences would be judged as unacceptable, which did not happen. This is a puzzle that needs to be addressed in the future. Perhaps speakers judged the sentences without taking into account the context. However, as there were only two sentences with the existential context, I will refrain from making any strong claim about the results.

(13) a. Joaquim's mother has a store called 'Just skirt.'
b. Nessa loja **se** vende saia.
In.this store sell:3SG skirt
'In this store one sells skirts.'

(14) a. Mariana's father has a store called 'Just dessert.'
b. **Se** vende doce.
sell:3SG dessert
'One sells desserts.'

6. Inferential Statistics

Until now I described the results from the sample I obtained. Now we want to make inferences about the population studied based on that sample. We should keep in mind, though, that our sample is small: with a larger sample, the inferences would be more solid. The hypotheses formulated at the beginning of the chapter were:

(i) In null generic impersonal constructions, the EPP needs to be checked by an element other than the null pronoun which values T's φ-features.

(ii) The generic reading of null impersonal sentences requires an overt marker of genericity.

Using the Anderson-Darling test, we see that the data do not follow a normal distribution: AD = 4.03, $p < 0.005$ for Test 1; AD = 5.11, $p < 0.005$ for Test 2. As the data do not follow a normal distribution, a non-parametric test, the Friedman Test, similar to the parametric Repeated Measure ANOVA, was used to detect differences across conditions within Test 1 and within Test 2. That is, one separate Friedman Test was run for each test.

For Test 1 (*without se*), when the conditions ADV, ADV+D, ADV+NEG+D, NEG+D, ADV+NEG, NEG and NoADV were compared, a significant difference was detected (chi-square = 18.52 $p < 0.01$, 2-tailed). A post hoc analysis was conducted to see in which conditions the difference was found. The results are reported in the box below. Conditions with the same letter are not significantly different:

Box 1 – Post hoc analysis: sentences without se (Test 1)

Post Hoc Analysis
Alpha: 0.05; DF Error: 60
t-Student: 2.00298
LSD: 16.37314

Condition	Median	Mean	Sum of ranks	Groups
ADV+D	5	4.77	67.0	a
ADV+NEG+D	5	4.59	61.0	a
NEG+D	5	4.50	60.0	a
NEG	5	4.36	42.0	b
ADV+NEG	4	3.64	36.0	b
NoADV	4	3.54	26.5	c
ADV	3	3.27	15.5	d

I will first discuss whether the results support the hypothesis in (ii). I will address hypothesis (i) after this discussion. As explained before, while the sentences in the conditions ADV+D, NEG+D and ADV+NEG+D had an overt marker of genericity (i.e., the necessity deontic modal), the sentences in ADV, NEG, ADV+NEG and NoADV did not have an overt marker of genericity, only a background favoring the generic reading. In Box 1, we see that the conditions that received a higher score (*group a*) are the ones with a modal in contrast to the ones without a modal (*group b, c and d*). This seems to point out that participants at least prefer sentences with an overt marker of genericity as opposite to sentences without it when they are presented with a background favoring the generic reading. The results do not suggest that the overt marker of genericity is *required* in generic sentences, only that it is preferred, as sentences without a modal were not judged as completely ungrammatical.

Notice that ADV and NoADV received a significant lower score than the other sentences without a modal (still, all sentences without a modal were judged as worse than sentences without it). Recall that there was only one generic sentence for each one of these conditions, while for the other conditions, there were two items. A balance in the number of sentences per condition could have rendered different results.

The condition NEG+D was important for testing the hypothesis (i), namely, that the EPP should be satisfied by some element other

than the null generic pronoun. The prediction is that sentences in this condition should be unacceptable, if there is an EPP feature that needs to be checked by elements such as locatives occurring in [Spec, TP]. In Box 1, we also see that the condition NEG+D was not significantly different from generic structures with an adverbial that could check the EPP. It is necessary to take into account, though, that there was only one condition in Test 1 that tested the role of the EPP. The sentences in the conditions NoADV and NEG did not have an element that could check the EPP, but they also did not have an overt marker of genericity, in such a way that the absence of an overt marker of genericity is enough to explain the low average score obtained for these two conditions.

For Test 2 (*with se*), a significant difference was detected when the conditions ADV, ADV+D, ADV+NEG+D, NEG+D, ADV+NEG, NEG and NoADV were compared (chi-square = 18.51, $p < 0.01$, 2-tailed). A post hoc analysis was run to detect which conditions differ from each other. The results of the post hoc test can be seen below in Box 2. Conditions that share the same letter are not significantly different:

Box 2 – Post hoc analysis: sentences with se (Test 2)

Post Hoc Analysis
Alpha: 0.05; DF Error: 60
t-Student: 2.00298
LSD: 9.382228

Condition	Median	Mean	Sum of ranks	Groups
NEG	5	4.72	55.5	a
ADV+NEG	5	4.59	52.5	a
ADV+NEG+D	5	3.27	52.0	a
NEG+D	4	3.68	52.0	a
ADV	3	3.27	35.5	b
NoADV	4	3.54	31.0	b
ADV+D	2	2.18	29.5	b

It was proposed in section 1 that *se* and the deontic modal function as overt markers of genericity. It seems, though, that nothing prevents *se* from co-occurring with a modal, given the fact that the conditions ADV+NEG+D and NEG+D (*with se*) are not significantly different

from the conditions with the highest score without a deontic modal (ADV+NEG and NEG (*with se*)).

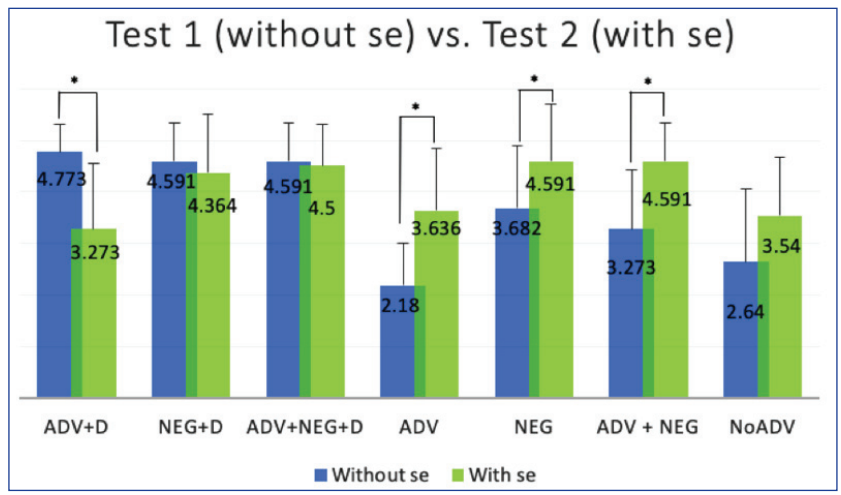
Participants were reluctant in accepting *tem que* (have to) with *se,* since the condition ADV+D was judged as bad as the condition NoADV and ADV (in which there was only one sentence per condition): it was pointed out above that *tem que* (have to) has a different interpretation than *não pode* (cannot) which was used for the conditions ADV+NEG+D and NEG+D. That might be the reason why participants in general rejected *tem que* with *se*. This result warrants further investigation to examine whether *tem que* in fact cannot co-occur with *se* and which properties are responsible for that.

The conditions ADV and NoADV in Test 2 are likely to have received a low average score compared to the sentences for other conditions for the same reason as in Test 1: the conditions ADV and NoADV had only one sentence each, while the other conditions had two sentences each. A balance in the number of sentences per condition could have rendered different results.

Notice that the score for the condition NEG and NEG+D were statistically identical to the conditions ADV+NEG, ADV+NEG+D and there is no adverb in the sentences for the conditions NEG and NEG+D that could check the EPP feature. The hypothesis (i) is not supported by the results of Test 2 and it is not supported by the results of Test 1 as well, as discussed above.

In Graph 5, we plotted the bar graphs for Test 1 (*without se*, in blue) and Test 2 (*with se,* in green) together. Examining Graph 5, a clearer pattern emerges and we are able to better analyze whether the results support the hypothesis (ii) or not (i.e., impersonal structures with the generic reading require an overt marker of genericity). Notice that all conditions without a modal, with the exception of NoADV, were judged as significantly better when *se* was included (Test 2) than when just a null subject was present (Test 1). These results suggest that the overt marker of genericity improves the availability of the generic reading. However, as sentences without an overt marker of genericity were not judged as completely ungrammatical, we can only say that BP-speakers *prefer* generic sentences with an overt marker of genericity.

Graph 5 – Average for each structure: sentences without se and with se



In order to support the hypothesis that sentences without an overt marker of genericity are not completely ungrammatical, I conducted the Friedman for Test 1 (*without se*) and Test 2 (*with se*) including now the ungrammatical filler. The prediction is that a significant difference should be found between the ungrammatical filler and all the other sentences, especially the ones without an overt marker of genericity. For Test 1 (*without se*), a significant difference was found between conditions (chi-square = 63.34, $p < 0.001$). The post hoc test detected a significant difference between the ungrammatical filler and all other conditions, with the filler receiving the lowest average score (1.18).

Although all the sentences in Test 2 (*with se*) had an overt marker of genericity, the Friedman test including the ungrammatical filler was also conducted to verify the grammaticality of the test sentences. For Test 1 (*with se*), a significant difference was found between conditions (chi-square = 3.17, $p < 0.001$). The post hoc test detected a significant difference between the fillers and all the other conditions, with the filler receiving the lowest average score (1.45).

Before proceeding in explaining the pattern seen in Graph 5, I describe how the statistical analysis comparing the two tests (Test 1 *vs* Test 2) was conducted. Pairwise Mann-Whitney tests were performed to compare each one of the conditions in Test 1 and Test 2. The Mann-

Whitney test was chosen because the samples are independent (different participants took Test 1 and Test 2). As only one comparison was performed by condition, the tests were performed without a correction. The table below shows the values of U and the p-values for each one of the comparisons (e.g., ADV (1) and ADV (2) stands for ADV (Test 1) and ADV (Test 2), respectively).

Table 3 – Values of U and p-values for Test 1 (*without se*) vs. Test 2 (*with se*)

Conditions	Values of U	P-values (2-tailed)
ADV (1) - ADV (2)	22	$p < 0.05$
NEG (1) - NEG (2)	24	$p < 0.05$
ADV+NEG (1) - ADV+NEG (2)	18.5	$p < 0.05$
ADV+NEG+D (1) - ADV+NEG+D (2)	55.55	$p = 0.76$
ADV+D (1) - ADV+D (2)	17	$p < 0.05$
NEG+D (1) - NEG+D (2)	59	$p = 0.94$
NoADV (1) - NoADV (2)	34	$p = 0.87$

I explain in detail the results for each one of the conditions below:

ADV: we expected that sentences *with se* would be judged as better than sentences *without se*, considering the hypothesis that the generic reading of impersonal sentences requires an overt marker of genericity in BP. As *se* is an overt marker of genericity, ADV is expected to be better *with se* than *without* it, as we would have no other overt marker of genericity in this kind of sentence. The test sentence *with se* rendered a higher score than the test sentences *without se* and the difference was statistically significant. The results do not support the hypothesis that the generic marker is *required*, but they suggest that speakers of BP *prefer* generic sentences with the overt marker of genericity.

Notice that the average score for this condition was low in both Test 1 and Test 2. It was pointed out that a factor that might have contributed to the low average score obtained for these conditions is that there was only one sentence for ADV (the other sentence was excluded from the statistical analysis because they had an existential

reading), while all the other conditions had two sentences each (with the exception of NoADV).

NEG: sentences *with se* in this condition should be judged as better than sentences *without se* for exactly the same reason as in ADV: adult speakers prefer generic sentences with an overt marker of genericity. As *se* is an overt marker of genericity, NEG is expected to be better *with se* than *without* it. For this condition, the test sentences *with se* were judged as significantly better than sentences *without se*, as predicted.

ADV+NEG: sentences *with se* in this condition would be judged as better than sentences *without se* for exactly the same reason as in ADV and NEG: adult speakers prefer generic structures with an overt marker of genericity than without it. As *se* is an overt marker of genericity, ADV+NEG is expected to be better with *se* than without it. For this condition, the test sentences *with se* were judged as significantly better than sentences *without se*, as predicted.

ADV+NEG+D: sentences *without se* in this condition should be judged as equally good as sentences *with se* if nothing in the language prevents double markers of genericity (*se* and the modal). It seems that there is nothing that prevents it, since the score was almost at ceiling in both Test 1 (*without se*, with no double marker of genericity) and Test 2 (*with se,* with a double marker of genericity). For this condition, no difference was found between the test sentences *with se* and *without se*, as predicted.

ADV+D: sentences *without se* in this condition should be judged as equally good as sentences *with se* for the same reason as in ADV+NEG+D. However, as explained before, the deontic modal *tem que* ('have to') was used in ADV+D, while the deontic modal used in all the other conditions was *não pode* ('cannot'). It was suggested that speakers do not favor the co-occurrence of *tem que* ('have to') with *se,* explaining why sentences *without se* were judged as better than sentences *with se.* This pattern clearly needs to be investigated further.

NEG+D: sentences *without se* in this condition should be judged as equally good as sentences *with se* if nothing in the language prevents double markers of genericity (*se* and the modal), as in ADV+NEG+D

and ADV+D. For this condition, no difference was found between the test sentences *with se* and *without se*, as predicted.

NoADV: the results for this condition were very similar to the results for the condition ADV and this condition had the same problem as ADV (i.e., just one sentence for this condition). Overall, sentences *with se* were judged as better than sentences *without se*, but this difference was not statistically significant. A larger sample and more sentences in this condition could have rendered significant results.

The average score for this condition was low in both Test 1 and Test 2. Two factors might be responsible for that: (i) there was only one sentence for NoADV and (ii) the possibility that in the beginning of a sentence, *se* has the tendency to appear as an enclitic rather than a proclitic.

7. Conclusions

This article reported the results of an online survey conducted with adult native speakers of BP. The purpose of the survey was to see if the following hypotheses were supported:

(i) In null generic impersonal constructions, the EPP needs to be checked by an element other than the null pronoun which values T's φ-features.

(ii) The generic reading of null impersonal sentences requires an overt marker of genericity.

Regarding (i), the results of the survey do not support the hypothesis that the EPP must be checked by elements such as adverbials in generic impersonal structures. The condition NEG+D (Test 1) and NEG (Test 2) were employed to check the role of the EPP: the prediction was that test sentences in these conditions would be unacceptable for the participants, because there is no element that can check the EPP in these structures. However, the average score for both conditions was almost at ceiling.

The hypothesis (ii) was shown to be too strong: the results only suggest that speakers *prefer* generic impersonal structures with an

overt marker of genericity, not that they *require* these markers. That speakers prefer generic sentences with an overt marker of genericity is supported by the following results: when comparing Test 1 (*without se*) with Test 2 (*with se*), we saw that in sentences without a modal (ADV, NEG, ADV+NEG and NoADV), the presence of an overt marker of genericity (*se*) in general improved the judgement of the test sentences. Also, in Test 1 (*without se*), sentences with the deontic necessity modal were judged as better than sentences without it.

Conflict of interests

The author declare they have no conflict of interest.

References

Bertolino, K. (2017). *Modals and impersonal null subjects in Brazilian Portuguese* (Master's thesis). University of Connecticut.

Bertolino, K. (2020) *An Experimental stury on the acquisition of impersonals in Brazilian Portuguese* (Doctoral dissertation). University of Connecticut.

Carvalho, J. (2018). Diferentes tipos de sujeitos nulos no português brasileiro. *ReVEL*, *16*(30), 78–106.

Carvalho, J. (2019). Incorporated subjects in existential impersonal sentences in Brazilian Portuguese. In P. Herbeck, B. Pöll, & A. C. Wolfsgruber (Eds.), *Semantic and syntactic aspects of impersonality*. Buske.

Cinque, G. (1988). On Si-Constructions and the Theory of Arb. *Linguistic Inquiry*, *19*(4), 521–581.

D'Alessandro, R. (2007). *Impersonal "si" constructions: Agreement and Interpretation* (Vol. 90). Mouton de Gruyter.

Ferreira, F. (2017). *O Crescimento do e-commerce no Brasil e seus desafios logísticos* (Monograph). Universidade Candido Menes.

Funari, P. (1999). Como se tornar arqueólogo no Brasil. *Revista USP*, *44*, 74–85.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases* (Doctoral dissertation). University of Massachusetts.

Holmberg, A. (2005). Is There a Little Pro? Evidence from Finnish. *Linguistic Inquiry*, *36*(4), 533–564. https://doi.org/10.1162/002438905774464322

Holmberg, A. (2010a). Null subject parameters. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan (Eds.), *Parametric Variation:*

Null Subjects in Minimalist Theory. Cambridge University Press. https://doi.org/10.1017/CBO9780511770784

Holmberg, A. (2010b). The null generic subject pronoun in Finnish: A case of incorporation in T. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan (Eds.), *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge University Press. https://doi.org/10.1017/CBO9780511770784

Holmberg, A., Nayudu, A., & Sheehan, M. (2009). Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, *63*(1), 59–97. https://doi.org/10.1111/j.1467-9582.2008.01154.x

Othero, G. A.; Miranda da Silva, S. (2020). A indeterminação do sujeito no português brasileiro: investigando algumas estratégias pronominais em orações finitas. *Muitas Vozes*, v. 9, n. 1.

Vargas, A. S. C. (2010). *Estratégias pronominais de indeterminação* (Master's thesis). Universidade do Rio de Janeiro, Rio de Janeiro.

Vargas, A. S. C. (2012). A evolução na representação das estratégias pronominais de indeterminação. In D. Mel (Eds.). *O sujeito em peças de teatro (1833-1992): estudos diacrônicos.* Parábola, 45-67.

Recebido em: 30.10.2020
Aprovado em: 13.07.2021

APPENDIX
Sentences and backgrounds used in the survey

The order in which the sentences appeared was randomized for each participant who took the survey.

ADV (generic)
Concentração para estudar é um desafio que persegue todos os estudantes. Na USP, há muitas bibliotecas disponíveis e é comum que os alunos estudem lá, procurando por lugares quietos para ler ou fazer um trabalho.
Concentration while studying is a challenge for every student. At USP, there are many libraries available and it is common that students go there to study, searching for a quiet place to read or work on assignments.

Nessa	universidade	(se) estuda	na	biblioteca.
In.this	university	study:3SG	in.the	library

'At this university one studies in the library.'

ADV (existential)
A mãe do Joaquim tem uma loja chamada 'Só saia'.
Joaquim's mother has a store called 'Only skirts'.

Nessa	loja	(se) vende		saia.
In.this	store	sell:3SG		skirt

'In this store someone sells skirts.'

NEG
Nos Estados Unidos, o bolo de cenoura é feito com uma cobertura de cream cheese, ao invés da cobertura tradicional de chocolate que a gente tem aqui. No Brasil é diferente.
In the US, carrot cakes are made with cream cheese frosting, instead of the traditional chocolate frosting we have here. In Brazil, things are different.

Não (se)	faz	bolo	de cenoura	desse	jeito.
Not	make:3SG	cake	of carrot	in.this	way

'One does not make carrot cake in this way.'

Se você era uma criança nos anos 90, você muito provavelmente tinha uma vizinha que vendia geladinho. Mas, as coisas mudaram.
If you were a child in the 90s, you probably had a neighbor who used to sell freeze pops. But then things changed.

Não (se) vende mais geladinho.
Not sell:3SG more freeze.pop
'One does not sell freeze pops anymore.'

NoADV (generic)
Além do clima frio fazer com que o banho diário não seja uma necessidade para os americanos, alguns acreditam que esse hábito prejudica a pele e os cabelos. No Brasil, é normal e higiênico tomar banhos diariamente.
The cold weather makes daily showers unnecessary for Americans. Besides, some of them believe that daily showers damage the skin and the hair. In Brazil, it is normal and hygienic to shower daily.

(Se) toma banho todo dia.
take:3SG shower every day
'One showers every day.'

NoADV (existential)
O pai da Mariana tem uma loja chamada 'Só doce'.
Mariana's father has a store called 'Only dessert'.
Mariana's father has a store called 'Only dessert'

(Se) vende doce.
sell:3SG dessert
'Someone sells desserts.'

ADV+NEG
Na Polônia, muitas famílias gostam de colher cogumelos nos bosques, que cobrem mais de 30% do país. Poloneses não acham que é perigoso comer cogumelos selvagens, mas isso não acontece em todas as culturas.
In Poland many families like to pick mushrooms in forests which cover more than 30% of the country. Polish people do not think it is dangerous to eat wild mushrooms, but that does not happen in every culture.

No Brasil não (se) colhe cogumelo em bosques.
In Brazil not pick:3SG mushroom in forests
'In Brazil one does not pick mushrooms in forests.'

Novelas brasileiras são famosas ao redor do mundo, mas essa popularidade não alcançou todos os países. Em 1995, a guerra da ex-Iugoslávia parou durante 1 semana para que o país assistisse os últimos capítulos da 'Escrava Isaura'! Mas, como dito, não é em todo lugar que as pessoas assistem novela.
Brazilian soap operas are famous around the world, but this popularity did not reach every country. In 1995, the Yugoslav War stopped for 1 week so that the

country could watch the last episodes of 'Slave Isaura'! But, as said before, it is not everywhere that people watch soap operas.

Nos Estados Unidos não (se) assiste novela.
In.the US not watch:3SG soap.operas
'In the US one does not watch soap operas.'

ADV+D
O Joaquim estuda em uma escola em que é obrigatório comer verdura na hora do almoço.
Joaquim studies in a school where it is mandatory to eat vegetables at lunchtime.

Nessa escola tem que (se) comer verdura na hora do almoço.
In.this school have:3SG that eat:INF vegetable at.the time of.the lunch
'In this school one has to eat vegetables at lunchtime.'

Na escola em que a Mariana estuda, tem uma regra em que todo dia os alunos precisam entregar a lição de casa. Quem não entrega a lição de casa, leva suspensão. É uma escola muito rígida.
In the school where Mariana studies, there is a rule which says that students have to turn in homework assignments every day. Who does not do that is suspended from school. This school is very strict.

Nessa escola (se) tem que entregar lição de casa.
In.this school have:3SG that turn.in:INF lesson of home
'In this school one has to turn in homework assignments.'

NEG+D
O abandono escolar é um fenômeno que tira da escola milhares de alunos que podem ser excluídos da sociedade e do mercado de trabalho.
Dropout is a phenomenon that removes several students from school who can be excluded from the society and job market.

Não (se) pode largar a escola.
Not can:3SG leave:INF the school
'One cannot drop out of school.'

De acordo com as leis brasileiras de trânsito, é proibido pedalar na calçada, a não ser que exista sinalização permitindo.
According to the Brazilian traffic regulations, it is forbidden to ride a bicycle on the sidewalk, unless there is a sign allowing it.

Não (se) pode andar de bicicleta na calçada.
Not can:3SG ride:INF of bicycle on.the sidewalk
'One cannot ride a bicycle on the sidewalk.'

ADV+NEG+D

Na escola em que a Mariana estuda, tem uma regra em que os alunos estão proibidos de trazer qualquer animal de estimação.
In the school where Mariana studies there is a rule in which students are forbidden to bring pets.

Nessa escola não (se) pode trazer animal de estimação.
In.this school not can:3SG bring:INF pet
'In this school one cannot bring pets.'

De maridos esquecidos o mundo está cheio, mas Samoa, na Oceania, está tentando melhorar isso. Lá é proibido esquecer o aniversário da esposa, e o marido deve responder à justiça e ainda indenizar a esposa se fizer isso.
The world is full of husbands with poor memory, but Samoa, a country in Oceania, is trying to change it! In Samoa a husband is forbidden to forget his wife's birthday. If he does, he is brought to justice and has to indemnify his wife for his actions.

Em Samoa não (se) pode esquecer o aniversário da esposa.
In Samoa not can:3SG forget:INF the birthday of.the wife
'In Samoa one cannot forget one's wife's birthday.'

FILLERS

O Joaquim está muito cansado depois de um dia na escola em que ele não estudou nada, mas conversou o tempo inteiro com os seus colegas. Quando ele chegou em casa, deitou em sua cama e dormiu tranquilamente por 12 horas.
Joaquim is very tired after a day at school in which he did not study at all, but talked the whole time with his classmates. When he got home, he lay on the bed and slept peacefully for 12 hours.

*A cama foi deitada pelo Joaquim.
The bed was laid.down by Joaquim
'The bed was laid down by Joaquim.'

A maioria das pessoas trabalha todos os dias de segunda à sexta, 8 horas por dia. Para manter uma vida saudável e ter tempo para o lazer e para a família…
Most people work every day from Monday through Friday, 8 hours a day. .In order to lead a healthy lifestyle and have time for leisure and family…

No fim de semana é melhor descansar.
On.the weekend is best rest:INF
'It is best to rest on weekends.'

Dentre as pragas que atacam o coco, destaca-se o ácaro da necrose. Essa é uma praga que anda causando muito prejuízo no Brasil. Não tem problema nenhum consumir a água ou a polpa do coco afetado pela praga, mas a aparência do coco dificulta a sua venda.
Among the pests that infest coconut plantations, the coconut mite is the main one. This is a pest that has been causing a lot of financial loss in Brazil. There is no problem consuming the water or the coconut pulp affected by the pest, but the appearance of the coconut makes it difficult to sell.

No Brasil, o ácaro da necrose dificulta a venda do coco.
In Brazil the mite of.the necrosis make.difficult the sales of.the coconut
'In Brazil the coconut mite makes it difficult to sell coconuts.'

A Mariana precisa doar suas roupas para o brechó. Mas, como as roupas estavam um pouco sujas, a Mariana decidiu colocar as roupas na máquina de lavar. Agora elas estão limpas!
Mariana needs to donate her clothes to the thrift store. But as the clothes were a little bit dirty, she decided to wash them in the washing machine. Now the clothes are clean!

? A roupa já lavou pra ser doada no brechó.
The clothes already washed to be donated in.the thrift store
'The clothes were already washed to be donated to the thrift store.'

O Bruno estava brincando na casa do Joaquim com um dinossauro de borracha que tinha dentes muito pontiagudos. Os dois meninos acabaram machucando o dedo.
Bruno was playing at Joaquim's house with a rubber dinosaur that had very sharp teeth. The boys ended up hurting their own fingers.

? O dinossauro do Joaquim$_i$ machucou o Bruno e ele$_i$ com os dentes.
The dinosaur of.the Joaquim hurt the Bruno and him with the teeth
'Joaquim$_i$'s dinosaur hurt Bruno and him$_i$ with his teeth.'