http://dx.doi.org/10.31513/linguistica.2022.v18n1a52459

Labeling small clauses[1]

Rotulando pequenas cláusulas

Marcelo Amorim Sibaldo[2]

ABSTRACT
There are several studies on the categorial status of *Small Clauses* (SCs) (STOWELL, 1981; MORO, 2000, and others), PrP (BOWERS, 1993), and PredP (BAILYN, 2001) etc.). Given that a theory of exocentric phrase structure requires Syntactic Objects to be labeled so that Full Interpretation can interpret it in CI, we ought to understand which label is given to SCs. The objective of this paper is to analyze data from Portuguese and Russian to argue that the system presented in two recent papers, *Problems of Projection, PoP,* (CHOMSKY, 2013, 2015), can explain the case and agreement morphology within SCs crosslinguistically. I propose that SCs can be labelled as ϕ in the PoP system when subject and predicate share these features. Otherwise, in case of mismatch between the morphological specifications of the subject-predicate set, the subject must be raised so that the predicate's case feature can be valued.
KEYWORDS: Small Clause. Case. ϕ-features. Label.

RESUMO
Existem vários estudos sobre o *status* categorial das Small Clauses (SCs) (STOWELL, 1981; MORO 2000, e outros), PrP (BOWERS, 1993), e PredP (BAILYN, 2001) etc.). Dado que uma teoria da estrutura de frases exocêntricas requer que os Objetos Sintáticos sejam rotulados para que o princípio de Interpretação Plena possa interpretá-los em no componente CI, devemos entender qual rótulo é dado às SCs. O objetivo deste artigo é analisar dados do português e do russo para argumentar que o sistema apresentado em dois artigos recentes, Problems of Projection, PoP, (CHOMSKY, 2013, 2015), pode explicar a morfologia de caso e concordância dentro de SCs de forma interlinguística. Proponho que as SCs podem ser rotuladas como ϕ no sistema PoP quando sujeito e predicado compartilham esses traços. De outro modo, em caso de incompatibilidade entre as especificações morfológicas do conjunto sujeito-predicado, o sujeito deve ser alçado para que o traço de Caso do predicado possa ser checado.
PALAVRAS-CHAVE: *Small Clause*. Caso. Traços-ϕ. Rótulo.

Introduction

Despite the intense debate around the topic, there seems to be consensual in the Generative

[1] A slightly modified version of this paper was published in Portuguese as Sibaldo (2020). Parts of this work were presented at I *Linguística Formal* (Federal University of Santa Catarina – UFSC) and at the Intermediary Summit of the Grammar Theory Work Group of ANPOLL (Catholic University of Rio de Janeiro – PUC-Rio). I would like to thank these events' organizers for their invite, especially Cilene Rodrigues, Cristina Figueiredo Silva, and Marco Antonio Martins. I would also like to thank the audiences of these conferences and Marcel den Dikken, Michelle Sheehan, Ian Roberts, Jairo Nunes, and Marco Antonio Martins, as their comments and suggestions help enhance this paper. Remaining errors are my sole responsibility. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and by UFPE.

[2] Professor of Portuguese at Universidade Federal de Pernambuco (UFPE), marcelo.sibaldo@ufpe.br, https://orcid.org/0000-0003-2119-8899.

Recebido em 16/05/2022 - Aceito em 15/08/2022

literature that sentences involving copulative verbs are generated at the base by what is known as a Small Clause (hereinafter denominated SC), i.e., a subject-predicate structure in which the latter is s-selected by the former.

However, within the theoretical apparatus of Generative Grammar, there are several questions pertaining to SCs "layout" which will be addressed here, such as: Do these elements have a head or do they have a "bare" configuration? Is there a label for SCs? If so, what would it be? These questions derive from the Government & Binding apparatus, within X-bar theory's endocentric model, and go through at the first moment of the Minimalist Program (MP), in Bare Phrase Structure Theory. They remain in a more recent version of MP (an exocentric model of syntactic objects). Chomsky published this version in *Problems of Projection* (PoP) in 2013 and its continuation in 2015 — *Problems of Projection: Extensions*, proposing that the simpler merge does not assign a label to its output.

This paper aims to descriptively show that case morphology in Russian and inflectional morphology (ϕ-features[3]) in Portuguese may bear a similar explanation in SC contexts and, theoretically, argue that an exocentric theory of phrasal structure labeling (cf. CHOMSKY, 2013 and 2015) may explain case and inflectional morphology in SCs, as proposed by Moro (2000 and 2006).

The remaining of paper is divided in 3 main sections. Section 1 introduces my understanding of Small Clauses and summarizes the proposals available in the literature to account for the configurations of these syntactic object. Section 2 addresses the role of syntactic object labels based on the PoP system (CHOMSKY, 2013, 2015), and section 3 puts forward my SC labeling proposal within an exocentric model for syntactic object. Closing the paper, some final considerations are presented.

1. About small clauses and agreement

As far as I know, the first researcher to use the term *Small Clause* in Generative literature was Williams (1975) in a paper called *Small Clause in English*. Despite using the term *Small Clause* in his title, the author did not explore the concept in the study.

The first propositions on Small Clause intended to discuss English sentences whose verbs end in *-ing* (WILLIAMS, 1975), like the ones in (1), as well as showing that all larger syntactic projections (S, NP, AP, VP etc.) may include a subject position (STOWELL, 1981, 1983), as seen in sentences in (2):

(1) a. The man **driving the bus** is Norton's best friend.
b. John decided to leave, **thinking the party was over**.
c. **John's evading his taxes** infuriates me.

(WILLIAMS, 1975, p. 249, exs. (1)-(3))

[3] For the purposes of this paper, I understand ϕ-features as related to gender, number, and person features. Refer to Harbour, Adger, and Béjar (2008) for studies more extended notions of the concept.

(2) a. Alexandra proved [$_{AP}$ the theory [false]].
b. I expect [$_{PP}$ that man [off my ship]].
c. Mary had [$_{VP}$ her brother [open the door]]

(STOWELL, 1983, p. 297, exs. (24b), (25a), (26a))

The main difference between Williams' and Stowell's SC theories lies in the fact that the latter considers SC as forming one single constituent, whereas the former — especially in *Against Small Clause* (WILLIAMS, 1983) — considers SCs as involving more than one constituent. Stowell's theory became known in the literature as *The SC Theory* — in it, AP, PP, and VP in (2) are SCs and hence form one single constituent. In its turn, William's theory became known as *The Predication Theory* — for its advocates, by definition, the subject is not contained in the same phrase as the predicate, hence the name *external argument* for subjects in general.

I will leave larger matters for other studies. For now, I shall adopt Stowell's view and assume that a Small Clause is nothing more than a *predication* of a predicate over its subject. Therefore, in examples under (3), we see SC instances — base of a copulative sentence — and, in examples under (4), we see what is referred to in literature as a complement SC. Both examples include an SC with a predicate selecting a subject.

(3) a. A Maria está cansada.
"Maria is tired$_{sing/fem}$."
b. Os meninos estão cansados.
"The boys are tired$_{pl/masc}$."

(4) a. O João considera a Maria uma boa aluna.
"João considers Mary a good student."
b. O João considera os meninos bons alunos.
"João considers the boys good students."

In addition to gender and number agreement between [A Maria] and [cansada/uma boa aluna] (3a) and (4a), on one end, and [Os meninos] and [cansados/bons alunos] (3b) and (4b), on the other end — I shall address this later in this paper —, another way of saying that the predicative (SC's predicate) predicates over subjects is that there are semantic restrictions in relation to the subject, as we can see below:

(5) a. *O piso está cansado.
"The floor is tired."
b. *O João considera a pedra uma boa aluna.
"João considers the rock a good student."

Thus, with respect to sentences (3) and (4), data in (5) is ill-formed because the predicate poses semantic restrictions to the copulatives' subjects. In other words, it is the predicate that semantically selects (predicates) over the subject.

Still, on the morphologically visible agreement between subject and SC's predicates, one can crosslinguistically observe that this is true in relation to both case and inflectional morphology, as seen in the examples below (where agreement is mandatory):

(6) a. [Ciceronem clarum] habent. *Latin*
Cicero$_{ACC}$ *famous*$_{ACC}$ *consider/hold*
"They consider Cicero famous"

(MATUSHANSKY, 2008)

b. Hún er kennari/ *kennara. *Icelandic*
she-NOM is teacher-NOM/ ACC*
"She is teacher."

(MALING & SPROUSE, 1995, p. 168)

c. *O Pedro considera [a Maria bonito]. *Portuguese*
the$_{masc}$ Pedro considers the$_{fem}$ Maria beautiful$_{masc}$

However, several data from different languages show the lack of agreement is also possible:

(7) a. Ja sčitaju ee lingvistkoj. *Russian*
I consider her$_{ACC}$ *linguist*$_{INSTR}$
"I consider her a linguist."

b. Mari pea-b Jaani targa-ks. *Finnish*
*Mari.*NOM *considers Jaani.***PART** *intelligent.***TRS**
"Mari considers Jaani intelligent."

(MATUSHANSKY, 2006)

As this study is focused on copulative structures, I would like to introduce an even more interesting matter — in Russian, one can observe an asymmetry with respect to case morphology, as the predicate can showcase both nominative and instrumental cases in some copulative structures:

(8) a. Puškin byl velikij poèt. *Russian*
Pushkin was great poet$_{NOM}$

b. Puškin byl velikim poètom.
Pushkin was great poet$_{INSTR}$
"Pushkin was a great poet."

(MATUSHANSKY, 2006)

Therefore, as far as we can observe, the copula can be omitted in present tense in copulative sentences in Russian. Nevertheless, it is worth noting that, without the copula, only the nominative case in predicate is allowed:

(9) a. Vera assistent.
Vera assistant$_{NOM}$
"Vera is an assistant."
b. *Vera assistentom.
Vera assistant$_{INSTR}$

An interesting point arises when we compare data on case morphology in Russian with data on inflectional morphology in Portuguese, as seen below:

(10) a. As meninas estão bonitas.
the$_{pl/fem}$ *girls* *are*$_{estar}$ *pretty*$_{pl/fem}$
b. As meninas são bonitas.
the$_{pl/fem}$ *girls* *are*$_{estar}$ *pretty*$_{pl/fem}$
"Girls are pretty."

(11) a. As meninas são o orgulho do pai.
the$_{pl/fem}$ *girls* *are*$_{ser}$ *the*$_{sing/masc}$ *pride* *of.the father*
b. *?As meninas estão o orgulho do pai.[4]
the$_{pl/fem}$ *girls* *are*$_{estar}$ *the*$_{sing/masc}$ *pride* *of.the father*
"Girls are the father's pride"

In light of these analyses, I shall now address how Generative Grammar showcases SC's configuration in a tree diagram.

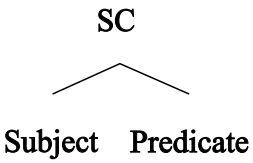
According to contemporary literature, copulative sentences *always* contain an SC in its base, whereas the opposite is not true, as an SC does not aways derive from a copular sentence, as observed in (12) below:

(12) a. A Maria é bonita.
b. [$_{TP}$ A Maria$_i$ [$_{T'}$ é [$_{SC}$ t_i bonita]]]
c. O João considera a Maria bonita.
d. O João considera [$_{SC}$ a Maria bonita]

[4] Some Portuguese speakers told me this sentence can work in specific contexts, maybe motivated by coercive effects (see MOLINA, 2008). Besides, to my understanding, these contexts are always taking *estar* (be) as an aspectual auxiliary and not as an "actual" copula, triggering the sentence "As meninas estão *sendo* o orgulho do pai" (The girls are being their father's pride), where the "actual" copula, *ser* (be), is observed.

In the current Generative Theory, there are two configurational propositions to accommodate data from sentences like (12a): (i) the bare Small Clause (symmetric); and (ii) the rich Small Clause (configurational). (i) is represented like (13) below:

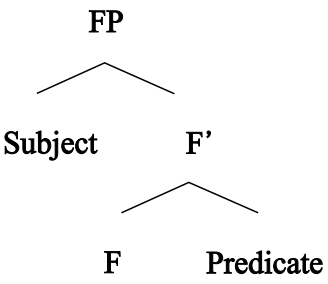
(13)



Moro (2000 and 2006) is one of the main authors in favor of (i). According to him, it is the symmetric structure in (13) that triggers syntactic movements, and not just morphological features, as Chomsky (2001) presumes. SC's structure in (13) is symmetric and thus violates Kayne's (1994) *Linear Correspondency Axiom* (LCA), which triggers subject or predicate raising to break down such symmetry.

Other authors, in their turn, presume a rich SC, as illustrated in (14) below:[5]

(14)



The structure in (14) is presumed by several authors due to both empirical data and theoretical matters (cf. BOWERS, 1993; CITKO, 2008; CONTRERAS, 1995; GUERÓN & HOEKSTRA, 1995; ADGER & RAMCHAND, 2003; HARVES, 2002; DEN DIKKEN, 2006; and many others.)[6]

A point worth mentioning is that, as it is placed, the structure in (13) poses a great problem for X-Bar Theory — which is essentially endocentric. This is because (13) does not have a head. Another point is that, as one can observe, neither (13) nor (14) carry any categorial status from SC. Is syntactic computation able to manipulate any categorial status in SCs? Some authors that presume a configurational status for SC has assigned to it different labels: AgrP (STOWELL, 1981 and 1983; MORO, 1988);[7] PrP (BOWERS, 1993); PredP (BAILYN, 2001); RP (DEN DIKKEN, 2006), among others.

Another question that arises is: Does the computational system really need that constituents have labels? Collins (2002, p. 42) states that the computational system does not require labels: "[...] the labels of phrasal categories (e.g. VP versus NP) are not needed in syntactic theory."

---

[5] Where F = any functional category.

[6] I shall not go into detail here due to space constraints.

[7] The idea behind this label *(Agr(eement))* has to do with the fact that, in some constructions, SC's subject and predicate must agree morphologically, as previously seen. Nevertheless, we have also seen that agreement is not aways triggered and, hence, this label stopped being used.

However, at a personal communication published by Fry (2013), Chomsky argues that labels are needed for interpretation:[8]

(15) The book John read

a. $[_{DP}$ [the book] $[_{CP}$ op C $[_{TP}$ John read <op>]]]

b. $[_{CP}$ C $[_{FocP}$ [the book] Foc $[_{TP}$ John read <the book>]]]

As we know, "the book" in (15) can be interpreted either as a DP that may be relativized, as in (15a), or as a focus, as in (15b). The system uses labels to provide this interpretation. I will come back to these arguments during my analysis of SC and of data from copulative sentences in Portuguese and Russian.

In the next section, I shall address the PoP system (CHOMSKY, 2013, 2015), an exocentric theory of syntactic objects that requires them to be labelled so that the *Principle of Full Interpretation* can be satisfied at the Conceptual-Intentional interface. I shall then apply such postulates to data from Russian and Portuguese with the aim to unify case and inflectional morphology, as it has been regularly done in the Generative theory.

2. Labeling syntactic objects

Chomsky (2013, 2015) puts forward an interesting proposal regarding theoretical linguistics' operation merge. The question is: When a Syntactic Object (hereinafter, SO) is merged to another, how does the system know which is the newly formed object's label? Chomsky proposes that SOs are built upon merge and a Labeling Algorithm, which assigns the "label" feature to labeled heads — hence, he proposes an exocentric theory for SOs (different than X-Bar's endocentricity).

Therefore, Chomsky presumes that labels are placed during derivation, and when two SOs are merged, no label is added to the newly created node. The operation responsible for the new node's label is what Chomsky (2015, p. 3) calls Labeling Algorithm (LA), "a special case of minimal search (as *Agree*), that fits within minimal computation." Labeling would then be a "label" feature to labeled heads. It is important to note that labeling does not produce a new category, but the SO can be labeled to be interpreted in the Conceptual-Intentional (CI) interface — otherwise, it violates the *Principle of Full Interpretation* (cf. EPSTEIN, KITAHARA & SEELY, 2014).

Chomsky (2015, p. 7) presents three contexts that trigger the three labeling possibilities:

(16) a. {H, βP}

b. {αP, βP}

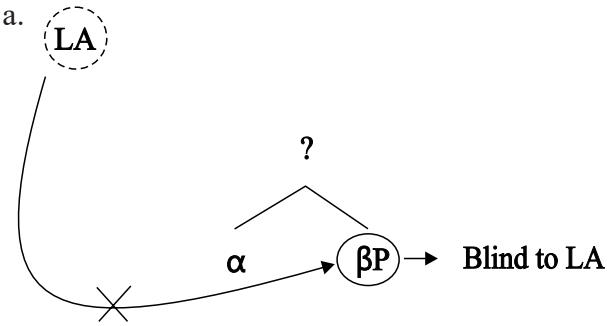
c. {α, β}

[8] In this example, according to Fry (2013), copies are indicated by chevrons: <, >. In other data mentioned in this paper, copies are indicated as usual by a t (for *trace*) or by a **contoured** font.

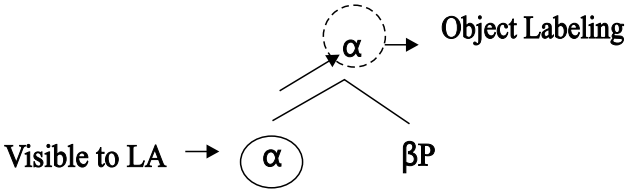
The contexts are: (16a) merge between a head and a phrase; (16b) merge of two phrases; and (16c) merge of two heads. For the purposes of this paper, I shall discuss (16a) and (16b).[9]

For Chomsky (2015, p. 7), labeling in (16a) — head-phrase merge — is "trivial": since the head is a lexical item, it makes available for the interfaces what matters, i.e., its categorial feature; however, the system cannot read a phrase in the same way as it is a *set* of other features. Therefore, to (16a), the label is always the head:

(17) a.



b.

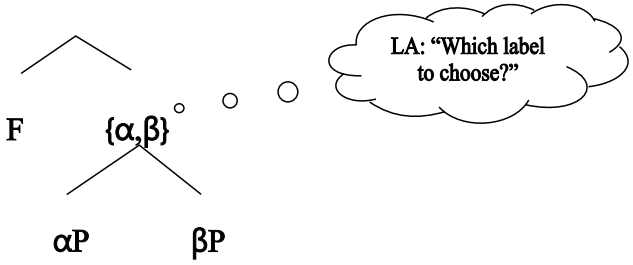


Consequently, Chomsky proposes that the LA seeks to label headless nodes so that they can be interpreted at the CI interface and, thus, meet the *Principles of Full Interpretation*. As it is "simpler," in the case of (17a), the node in question is labeled as a head, as it can be more quickly identified by the LA. This way, the system does not see any ambiguity in this labeling.

Nevertheless, the labeling process seen in (16b) — labeling of two phrases — gives rise to two possibilities, since the system can see both α and β sets.

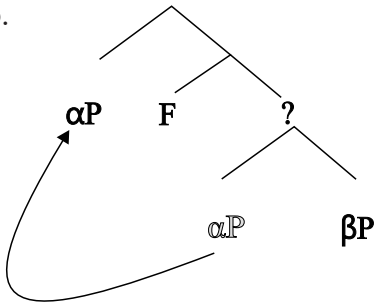
In the first scenario, Chomsky proposes that there is an SO modification, in the sense that one of the SOs must move so the LA is able to see the OS that remained *in situ* and, consequently, label the node created through this OS merge. Labeling takes place in a similar fashion to the examples below:

(18) a.

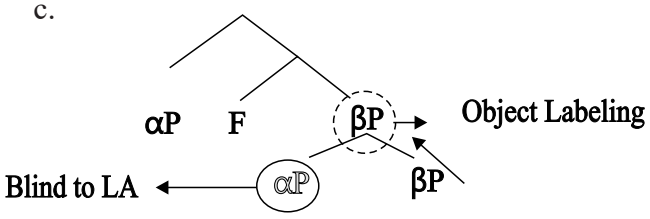


[9] For the case of two-heads merge, I refer to Saito (2013).

b.



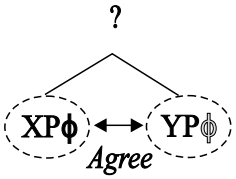
c.



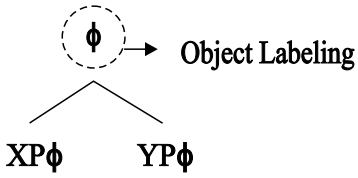
As seen in (18a), the algorithm does not understand which label to give to the newly created node — if none of the phrases move, there is no labeling and, consequently, derivation fails. (18b) shows that one of the (αP) sets is raised — the LA can only see the *in situ* phrase (βP), whose features are spread through the node and, then, LA labels the lowest node as βP (cf. 18c). In this scenario, raising one of the phrases is mandatory and must take place before labeling.

In the second scenario, in order to label the resulting node from the merge of the two phrases, Chomsky makes the following proposition: If the two SOs can share a relevant feature,[10] then there is no movement needed and the newly created node's label must be ϕ to indicate agreement. Refer to the step by step below:

(19) a.



b.



In (19a), upon the merge of two SOs (in this case, two phrases), the system tries to find a way of labeling the newly created node. These two SOs share common features, represented in (19) by ϕ. As in Chomsky's (2001) system, while one of the phrases — let us say XP — carries an interpretable version of this ϕ-feature set, the other phrase — YP — carries the non-interpretable version. By matching these two ϕ-feature sets, *Agree* happens and the algorithm labels the node as ϕ, exactly as in (19b). Now, this SO can be interpreted at the CI interface, meeting the *Principle of Full Interpretation*.

This was a brief overview of the theoretical framework I used to analyze SCs in Russian and Portuguese. Now, I shall go on to present the data.

3. The SC projection issue

To begin this section, let us consider the sentence below:

(20) João considera [$_{SC}$ a Maria muito bonita].

João considers the$_{sing/fem}$ *Maria very beautiful*$_{sing/fem}$.

"João considers Maria very beautiful."

[10] To Chomsky (2015), the phrases must share relevant features for labeling to take place. In addition to ϕ, he proposes that the same happens to Q-features shared in interrogative sentences, for example.

How does the system label the SC in (20) in its derivation? In other words, which would be the next step after (21)?

(21) a. Merge(DP,AP) → {$_{DP}$ *a Maria,* $_{AP}$ *muito bonita*}

b. {$_{?}$ {$_{DP}$ *a Maria,* $_{AP}$ *muito bonita*}}

As previously mentioned, at an SO {XP, YP}, LA can see both categorial features in (21), namely D and A, respectively. Two possibilities arise: (I) either there is a modification of the SO and one of the elements moves so that the LA can "see" the *in-situ* object's label and then label the newly created node; or (ii) D and A share a relevant feature (Q or ϕ) and this feature can thus work as a label for the SC's node.

As a result, based on the PoP's idea of labeling through identifying ϕ-features that Chomsky uses for other syntactic contexts — such as in (22) and (23) —, it is reasonable to say that (24) takes place in SCs, since DP and AP share these features.[11]

(22) [$_{?}$ $DP_{i\phi}$ [$TP_{u\phi}$]] → [$_{\phi}$ $DP_{i\phi}$ [$TP_{u\phi}$]]

(23) [$_{?}$ $DP_{i\phi}$ [$AP_{u\phi}$]] → [$_{\phi}$ $DP_{i\phi}$ [$AP_{u\phi}$]]

(24) a. João considera *[$_{?}$ [$_{DP}$ a Maria] [$_{AP}$ muito bonita]]*

b. João considera *[$_{\phi}$ [$_{DP}$ a Maria] [$_{AP}$ muito bonita]]*

Note that, overall, this idea resembles the one raised by Stowell (1981, 1983) and Moro (1988) in the 1980s, which postulates that AgrP is the label for SCs. The proposition of ϕ postulation seems to contemplate the Portuguese facts, which do now allow for non-compatibility of ϕ-feature in the context of complement SC (cf. (25a)), as well as the Latin data presented above — in relation to case morphology (cf. (25b)) — and the data from copulative sentences in Icelandic (cf. (25c)):

(25) a. **O Pedro considera [$_{SC}$ a Maria muito bonito].* *Portuguese*
The Pedro considers the$_{sing/fem}$ *Maria very beautiful*$_{sing/masc}$
"Pedro considers Maria very beautiful."

b. Ciceronem clarum] habent. *Latin*
Cicero$_{ACC}$ *famous*$_{ACC}$ *consider/hold*
"They consider Cicero famous"

(MATUSHANSKY, 2008)

[11] Letters "u" and "i" followed by subscribed ϕ refer, respectively, to uninterpretability and interpretability of these features, which makes both phrases (DP and AP) able to be in an *Agree* relation and agree between themselves (see CHOMSKY, 2001).

c. Hún er kennari/ *kennara. *Icelandic*
she-NOM is teacher-NOM/ ACC*
"She is teacher."

(MALING & SPROUSE, 1995, p. 168)

However, as seen in the previous section, the proposition of labeling *all* SCs as ϕ does not account for several other pieces of data in which feature incompatibility is observed. I shall now focus on Portuguese and Russian data to address this issue.

I will start by exploring Russian and copulative sentence data. As previously shown, one can observe the following asymmetry in Russian:

(26) a. Puškin byl velikij poèt. *Russian*
Pushkin was great poet$_{NOM}$
b. Puškin byl velikim poètom.
Pushkin was great poet$_{INSTR}$
"Pushkin was a great poet."

(MATUSHANSKY, 2006)

(27) a. Vera assistent.
Vera assistant$_{NOM}$
"Vera is an assistant."
b. *Vera assistentom.
Vera assistant$_{INSTR}$

(MATUSHANSKY, 2008)

As previously discussed, both nominative and instrumental cases can be in the predicate (cf. (26)) — however, when the copula is absent, only the nominative case is licensed (cf. (27)). It is worth noting that Russian speakers realize a difference in interpretation whether the predicate is marked with nominative case or with instrumental case. Let us see the data below:

(28) a. Ivan byl xrabryj soldat. *Russian*
Ivan was brave.$_{NOM}$ *soldier.*$_{NOM}$
"Ivan was a brave soldier."
b. Ivan byl xrabrym soldatom.
Ivan was brave.$_{INSTR}$ *soldier.*$_{INSTR}$
"Ivan was a brave soldier."

(PERELTSVAIG, 2001, p. 98)

For Pereltsvaig (2001, p. 98), meaning is different in the pair illustrated in (28): While the nominative marking in (28a) assigns a permanent interpretation (at individual level) to the predicate, the instrumental marking in (28b) assigns a temporal interpretation (at stage level).

It is worth noting that this distinction is widely known in the literature in relation to the type of copula in Portuguese, *ser* or *estar*, as we can see below:

(29) a. As meninas são bonitas.
"The girls are$_{SER}$ pretty."

b. As meninas estão bonitas.
"The girls are$_{ESTAR}$ pretty."

Similarly to Russian data in (28), in Portuguese, the verb *ser* in (29a) triggers an interpretation of permanent state, i.e., the girls are naturally beautiful; while in (29b), the verb *estar* triggers the interpretation of a transient state.

It is interesting to note that, in Portuguese, an asymmetry pretty similar to the one found in (27) from Russian can be observed:

(30) a. [$_{gen:\ fem/\ num:\ pl}$ As meninas] *são* [$_{gen:\ masc/\ num:\ sg}$ o orgulho do pai].

b. *?[$_{gen:\ fem/\ num:\ pl}$ As meninas] *estão* [$_{gen:\ masc/\ num:\ sg}$ o orgulho do pai].

Broadly speaking, it seems that the gender and number features in the subject of the SC can be in disagreement with the verb *ser*, while agreement is mostly mandatory in the case of the verb *estar*. Refer to the example in (30) — it shows equative sentences, in which (30a) is a well-formed sentence while (30b), which presents a gender mismatch between subject and SC predicate, is not.

In addition, as pinpointed by Foltran and Rodrigues (2013), Portuguese allows one single copula with *ser*, which became known in literature as a "pancake sentence" — predicational copulative sentences (leveraging typology from Higgins 1976) that do not show gender and/or number agreement between subject and predicate; on the other hand, *estar* triggers agreement in Portuguese, as seen in the examples below:

(31) a. Alunas que bebem é chato.
students$_{FEM/PL}$ *who drink*$_{PL}$ *is annoying*$_{MASC/SING}$
"The situation of students drinking is annoying."

b. Crianças pequenas é divertido.
children$_{FEM}$ little$_{FEM/PL}$ is$_{SING}$ fun$_{MASC/SING}$
"Some activitiy involving little children is fun."

(FOLTRAN e RODRIGUES, 2013, p. 270)

(32) a'. *Alunas que bebem está chato.
students who drink ESTÁ annoying.
b'. *Crianças está divertido.
children ESTÁ fun.

(FOLTRAN e RODRIGUES, 2013, p. 287)

Portuguese also allows for predicative clitics, but then again, solely in the case of verb *ser*:

(33) a. A Maria o é.
the$_{sing/fem}$ Maria it is$_{SER}$
"Maria is it."
b. *A Maria o está.
the$_{sing/fem}$ Maria it is$_{ESTAR}$

Another important question now is reflecting upon the differences between *ser* and *estar* in Portuguese so we can understand the observed asymmetries. A likely indication may come from Gallego and Uriagereka (2016) — the authors go back to Benveniste (1966) and many others (URIAGEREKA, 2001; ZAGONA, 2008; BRUCART, 2010) to suggest that *ser*'s selection seems to be more basic than *estar*'s (p. 124). In other words: *estar* seems to be the result of the incorporation of copula *ser* plus another element (maybe a preposition, in the authors' view), as seen below:

(34) *Estar* = *ser* + X

I shall bring two pieces of empirical evidence that Gallego and Uriagereka (2016, p. 129) discuss and that can be relevant for my analysis. The first case is related to locatives formed with *estar*:

(35) a. Is John *(there)?
b. O João está (aí)?
the$_{masc/sing}$ João is$_{ESTAR}$ there
c. A festa é *(aí)?
the$_{masc/sing}$ party is$_{SER}$ there

Comparing Portuguese and English (which has one single copula, *be*), one can observe that, in locative structures, the use of "there" is mandatory in English while it is optional in Portuguese. This seems to imply that, in Portuguese, *estar* carries an embedded locative preposition and thus does not require the use of another. Also note that, using a sentence with a SER verb, the presence of the locative is obligatory (cf. (35c)).

(36) showcases a second data set that corroborates my proposal, also used by Gallego and Uriagereka (2016, p. 130):

(36) a. A dívida está sendo negociada. ESTAR >> SER
The debt is$_{ESTAR}$ *being*$_{SER}$ *negotiated*

b. *A dívida é estando negociada. SER >> ESTAR
The debt is$_{SER}$ *being*$_{ESTAR}$ *negotiated*

"The debt is being negotiated."

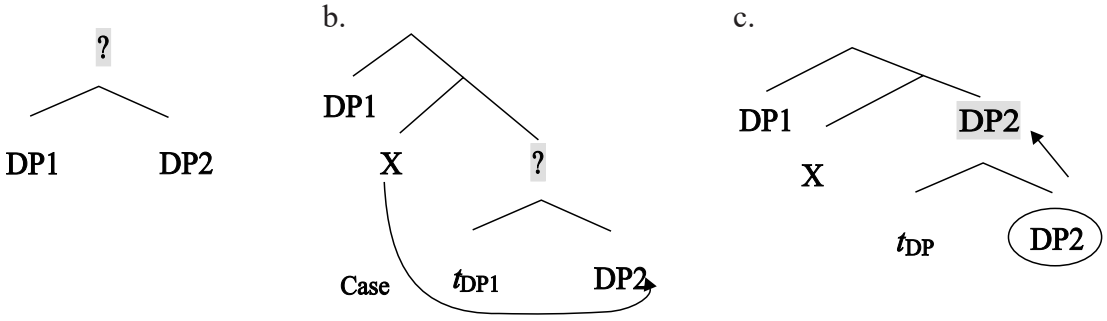
The data suggest that *estar* is derived at a higher position than *ser*.

Having that said, I shall address how Chomsky's (2013, 2015) labeling theory helps us understand what happens in relation to inflectional and case morphology in natural languages.

- Scenario 1: DPs that do not share features

Where features are incompatible, i.e., where SC's DPs do not agree — as in sentences with *ser* and in Chomsky's previously mentioned theory — one of the DPs must be raised. This leaves an element X (likely a preposition) that allows to verify the remaining DP (cf. (37b)). Since one of the DPs was raised, the LA can only see the label of the remaining DP; the SC's node is labeled as the phrase that remained *in situ* (cf. (37c)).

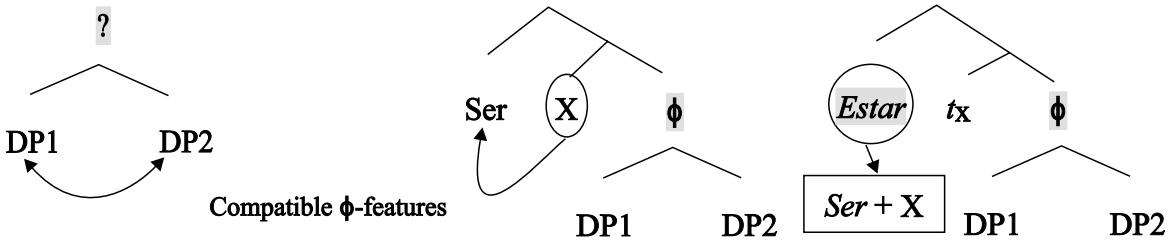
(37) a. b. c.



- Scenario 2: DPs that share features

When SC's DPs share relevant features (Q or ϕ, according to Chomsky), *Agree* happens and, consequently, the LA labels the SC node as ϕ (cf. (38b)). Since the element with prepositional value X was not used to verify Case (since *Agree* handles this), X is incorporated to copula *ser*, deriving *estar*.

(38) a b. c.



These two scenarios explain, for instance, why (27b) — repeated below as (39b) — is agrammatical:

(39) a. Vera assistent.
Vera assistant$_{NOM}$
"Vera is an assistant."
b. *Vera assistentom.
Vera assistant$_{INSTR}$

(MATUSHANSKY, 2008)

According to my analysis, (39a) is grammatical because the elements share the same case morphology, and labeling with ϕ happens due to *Agree*. (39b) is ungrammatical because subject and predicate do not share case morpheme and, hence, there is no feature compatibility. Since neither subject nor predicate are raised and both remain internalized within SC, labeling does not take place and the derivation crashes in the CI interface.

The system devised above predicts that, should any element intervene between subject and predicate of a copulative sentence, the sentence is grammatical even in the absence of any compatible case morphology. This happens because one of the DPs could have been raised from SC and, consequently, the system could have labeled the SC node with the DP that remained *in situ*. This is what happens when one adds a locative between subject and predicate, with distinct case morphemes, at a copulative in Russian:

(40) Saša **zdes'** studentom
*Sasha.**NOM** **here** student.**INSTR***
"Sasha is here as a student."

(BAILYN & EDWARD,1991, p. 121)

According to the system I proposed, since case morphemes of SC's subject and predicate in (40) are incompatible, the subject is raised so that the LA can label SC with the *in-situ* phrase and, thus, it is a well-formed sentence.

Let us get back to Portuguese: one can explain why *estar* can also select an NP as predicate. However, since it does not trigger subject agreement and the prepositional element has already been incorporated into the copula to form *estar*, a preposition must surface to verify the Case of the *in-situ* nominal. This can be observed in (41), where preposition *de* is realized to mark the Case of *babá* "babysitter".

(41) a. *João está babá.
João is$_{ESTAR}$ *babá*
b. João está **de** babá.
João is$_{ESTAR}$ ***of*** *babá*
"João is serving as a babysitter."

Lastly, I would like to show independent evidence from Portuguese for the system presented here. If I am correct, lack of agreement is accepted in NPs with one adjectival modifier agreeing with the head (as it is the case in 42a and 42b) provided that a preposition intervenes. Please refer to (42a) and (42b):

(42) a. [$_{NP}$ Amor materno] (**amor materna*) a'. [$_{NP}$ Amor de mãe]
[G: m, N: sg] [G: m, N: sg] [G: m, N: sg] [G: f, N: sg]
love motherly.$_{masc}$ *love motherly.$_{fem}$* *love of mother*

b. [$_{NP}$ Problemas estomacais] (**problemas estômago*) b'. [$_{NP}$ Problemas de estômago]
[G: f, N: pl] [G: f, N: pl] [G: f, N: pl] [G: m, N: sg]
problems stomach.pl *problems stomach.sing* *problems of stomach.sing*

As seen in the above data, in cases known by traditional grammar as "adjectival locutions," N and A share gender and number features when there is no preposition between name and adjective (such as in "amor materno" - maternal love - and "problemas estomacais" - stomach problems). Otherwise, a preposition is added, and non-agreement is likely to take place (cf. (42a'), where there is no gender agreement, and in (42b'), where there is neither gender nor number agreement.) According to my proposal, the latter is interpreted as a system's indication of lack of agreement between this phrase's elements. As a result, a preposition is added to verify the Case of the *in-situ* element. Unlike other proposals, I presume it is not the preposition that "blocks" agreement, as implied by previous frameworks — the preposition serves as a "warning" that the system needed to intervene to trigger a well-formed construction. This is necessary to ensure SOs are properly labeled and interpreted in the CI interface.

4. Final considerations

In this paper, I sought to argue in favor of an exocentric theory of syntactic structure labeling (CHOMSKY 2013, 2015) — in which *Merge* does not add labels — in order to unify the analysis of SC's case and inflectional morphology in Russian and Portuguese, as carried out by previous theoretical frameworks. This seems to be a *desideratum* from the theory since at least Chomsky (2000), who unifies Case-Agreement as a set.

Furthermore, I argued that, unlike in previous frameworks, prepositions can be seen as a system's indication of feature incompatibility between SOs.

References

ADGER, D.; RAMCHAND, G. Predication and Equation. *Linguistic Inquiry*, v. 34, n. 3, pp. 325-59, 2003.

BAILYN, J. The Syntax of Slavic Predicate Case. *ZAS Occasional Papers in Linguistics*, Berlin, 2001. pp. 1-26.

BAILYN, John; RUBIN, Edward J. The unification of Instrumental case assignment in Russian. *In*: *Cornell Working Papers in Linguistics*, vol. 9, ed. by A. Toribio and Wayne Harbert, 99-126. Ithaca, New York: Department of Modern Languages and Linguistics, Cornell University, 2001.

BENVENISTE, Émile. "Ser" e "Ter" nas suas Funções Lingüísticas. *In*: BENVENISTE, Émile. *Problemas de Lingüística Geral*. São Paulo: Companhia Editora Nacional, 1976. pp. 204-27.

BOWERS, John. The syntax of predication. *Linguistic Inquiry* v. 24, pp. 591-656, 1993.

BRUCART, J. M. La alternancia ser/estar y las construcciones atributivas de localización, in A. Avellana (ed.) *Actas del V Encuentro de Gramática Generativa.* Maestría en Lingüística, Universidad Nacional del Comahue, 2010. pp. 115-52.

CHOMSKY, N. Minimalist Inquires: The Framework. *In*: MARTIN, Robert; MICHAEL, David; URIAGEREKA, Juan. (eds.). *Step-by-Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: The MIT Press, 2000. pp. 89-155.

CHOMSKY, N. Derivation by Phase. *In*: KENSTOWICZ, M. (ed.). *Ken Hale*: A Life in Language. Cambridge, Mass: MIT Press, 2001. pp. 1-52.

CHOMSKY, Noam. Problems of projection. *Lingua* v. 130, pp. 33-49, 2013.

CHOMSKY, Noam. Problems of projection: Extensions. *In*: Di DOMENICO, Elisa; HAMANN, Cornelia; MATTEINI, Simona (eds.) *Structures, Strategies and Beyond*, 2015, pp. 1-16.

CITKO, Barbara. Small Clauses Reconsidered: Not so Small and Not All Alike. *Lingua*, 118, pp. 261-95, 2008.

COLLINS, C. Eliminating Labels. *In*: EPSTEIN, S. D.; SEELY, T. D. (eds.). *Derivation and Explanation in the Minimalist Program*. Malden, Mass.: Blackwell, 2002. pp. 42-64.

CONTRERAS, H. Small Clauses and Complex Predicates. *In*: CARDINALETTI, A.; GUASTI, M. T. (eds.). *Syntax and Semantics*: Small Clauses. v. 28. California: Academic Press, 1995. pp. 135-52.

DEN DIKKEN, Marcel. *Relators and linkers*: The syntax of predication, Predicate Inversion, and copulas. Cambridge, Massachusetts: MIT Press, 2006.

EPSTEIN, Samuel D.; KITAHARA, Hisatsugu; SEELY, Daniel. Labeling by minimal search: Implications for successive cyclic A-movement and the elimination of the postulate "phase". *Linguistic Inquiry,* v. 45, pp. 463-81, 2014.

FOLTRAN, Maria José D. G.; RODRIGUES, Patrícia Araújo. On Denoting Abstract Entities. *Revista da ABRALIN*, v. 12, pp. 269-91, 2013.

FRY, Brandon. Some notes on Chomsky 2013. Ms. University of Ottawa. Handout presented at Syntax-Semantics Reading Group, 2013.

GALLEGO, Ángel J.; URIEGEREKA, Juan. Estar = Ser + X. *Borealis*. *An International Journal of Hispanic Linguistics*, 5/11, 2016. pp. 123-56.

GALVES, C. *Ensaios sobre as gramáticas do português.* Campinas: Ed. Unicamp, 2001.

GUÉRON, J.; HOEKSTRA, T. The Temporal Interpretation of Predication. *In*: CARDINALETTI, A.; GUASTI, M. T. (eds.). *Syntax and Semantics*: Small Clauses. v. 28. California: Academic Press, 1995. pp. 77-107.

HARBOUR, D.; ADGER, D.; BÉJAR, S. *Phi-Theory*. Oxford: Oxford, 2008.

HARVES, S. Where Have All the Phases Gone? (Non-)Defective Categories and Case Alternations in Russian. *In*: TOMAN, J. (ed.). *Formal Approaches to Slavic Linguistics*: The Second Ann Arbor Meeting. Ann Arbor: Michigan Slavic Publications, 2002. pp. 97-118.

HIGGINS, F. R. *The Pseudo-Cleft Construction in English*. PhD Dissertation: Indiana University Linguistics Club, 1976.

KAYNE, R. S. *The Antisymmetry of Syntax*. Cambridge, Mass.: The MIT Press, 1994.

MALING, Joan; SPROUSE, Rex A. Structural case, specifier-head relations, and the case of predicate NPs. *In*: HAIDER, Hubert; OLSEN, Susan; VKNER, Sten (eds.). *Studies in Comparative Germanic Syntax*. Dordrecht: Kluwer, 1995, pp. 167-86.

MATUSHANSKY, Ora. Predication: a case study. *In*: F. Marušič and R. Žaucer, eds., Studies in Formal Slavic Linguistics. *Contributions from Formal Description of Slavic Languages 6.5*. Frankfurt am Main: Peter Lang, 2006, pp. 213-39.

MATUSHANSKY, Ora. Some cases of Russian, *Formal Description of Slavic Languages* (FDSL) 7.5. Moscow, December, 6-8, 2008.

MOLINA, S. G. Sobre las diferencias entre ser y estar. El tipo de predicado y el tipo de sujeto. *Redele*, n. 13, jun. 2008.

MORO, Andrea. Per una teoria unificata delle frasi copulari. *Rivista di Grammatica Generativa*, v. 13, 1988. pp. 81-110.

MORO, Andrea. *Dynamic Antisymmetry*, Cambridge, Massachusetts: MIT Press, 2000.

MORO, Andrea. *Some Notes on Unstable Structures*. Ms., Università San Raffaele di Milano, 2006.

PERELTSVAIG, Asya. *Copular sentences in Russian*. Dordrecht: Springer, 2007.

SAITO, M. Case and labeling in a language without ϕ-feature agreement. Ms. Nanzan University, 2013.

SIBALDO, Marcelo A. Small Clause é PoP. *Fórum Linguístico*, v. 17, pp. 4673-88, 2020.

STOWELL, T. *Origins of Phrase Structure*. Ph.D. Dissertation: MIT, 1981.

STOWELL, T. Subjects across Categories. *The Linguistic Review*, n. 2., 1983. pp. 285-312.

URIAGEREKA, J. *"Adjectival Clues,"* keynote speech at Acquisition of Spanish & Portugues/Hispanic Linguistics Symposium, University of Illinois at Urbana-Champaign (USA), October, 2001. pp. 11-14.

WILLIAMS, Edwin. Small Clauses in English. *In*: KIMBALL, J. *Syntax and Semantics*, v. 4. New York: Academic Press, 1975. pp. 249-73.

WILLIAMS, Edwin. Against Small Clauses. *Linguistic Inquiry*, 14, pp. 287-308, 1983.

ZAGONA, K. Ser and estar differ in both structure and aspectual features. *Plenary talk given at Chronos 8.* International Conference on Tense, Aspect, Mood, and Modality, University of Texas at Austin, 2008.