

Argument Coreference Between Content Verbs and Predicative Nouns in Brazilian Portuguese: A Pilot Corpus Study

Correferência entre argumentos de verbos plenos e nomes predicativos em português brasileiro: um estudo piloto baseado em corpus

Ryan Marçal Saldanha Magaña
Martinez

Universidade Federal de São Carlos
(UFSCAR) | São Carlos | SP | BR
ryan.saldanha.martinez@gmail.com
<https://orcid.org/0000-0002-2197-4352>

Abstract: The literature on complement clauses and support/light verb constructions provides most of the available information on coreference between arguments of matrix and subordinate syntactic predicates. However, this phenomenon is largely unexplored when it comes to verb phrases consisting of a content verb and a predicative noun in object position, such as *declarar paixão* (“declare passion”), *cumprir a promessa* (“fulfill the promise”), *responder perguntas* (“answer questions”), or *resistir à tentação* (“resist temptation”). The notions of obligatory and non-obligatory coreference do not fully explain this phenomenon, since non-coreferential usage of these phrases leads to different types of coercion, which are relevant properties to classify these phrases. This paper proposes a pilot study of verb phrases of this kind, extracting them from a Brazilian Portuguese newspaper corpus aiming at their syntactic-semantic classification. This procedure revealed a total of 75 such verb phrases. Based on semantic coercion and distributional properties, four major classes divided into seven subclasses are proposed to cover two-thirds of the data. There is also a fifth class with 17 items in which coreference appears to be a contextual property, i.e., unrelated to lexical semantics, and eight *hapax legomena*. The results provide a new perspective on this understudied topic, identifying also irregular aspects deserving further studies.

Keywords: predicative noun; coreference; support verb; light verb; coercion.



Resumo: A literatura sobre orações completivas e construções com verbo-suporte ou verbo leve provê a maior parte das informações disponíveis acerca da correferência entre argumentos de predicados sintáticos matriz e subordinados. Entretanto, esse fenômeno mantém-se amplamente inexplorado quando se trata de grupos verbais compostos de um verbo pleno e um nome predicativo em posição de objeto, como “declarar paixão”, “cumprir promessa”, “responder perguntas” ou “resistir à tentação”. As noções de correferência obrigatória ou não obrigatória não explicam completamente esses fenômenos, uma vez que o uso não correferencial desses grupos verbais leva a diferentes tipos de coerção, propriedades relevantes para classificá-los. Este artigo propõe um estudo piloto de tais grupos verbais, extraindo-os de um corpus jornalístico do português brasileiro com o objetivo de realizar sua classificação sintático-semântica. Tais procedimentos revelaram um total de 75 grupos verbais. Com base em coerção semântica e propriedades distribucionais, propõe-se que quatro grandes classes divisíveis em sete subclasses cobrem dois terços dos dados. Há ainda uma quinta subclasse de 17 itens que parece ter a correferência como uma propriedade contextual, isto é, não relacionada à semântica lexical, e oito *hapax legomena*. Esses resultados trazem uma nova perspectiva a um tópico pouco estudado, identificando também aspectos irregulares a serem explorados em maior profundidade em estudos subsequentes.

Palavras-chave: nome predicativo; correferência; verbo-suporte; verbo leve; coerção.

1 Introduction

Understanding how syntactic and/or semantic predicates interact in complex sentences has been challenging for linguistic investigations which seek to bridge lexical and phrase semantics. To what extent argument coreference, one of the most studied such phenomena, relates to semantics is a debated issue in the literature on complement clauses (e.g., Gross, 1975) and

support or light¹ verb constructions (Gross, 1981, 1998; Butt, 2010; Barros, 2014; Santos, 2015; Fotopoulou; Laporte; Nakamura, 2021; Calcia, 2022, Rassi, 2023, among others). The latter can be exemplified by constructions such as “give a kiss”, “take a look”, “have a dream”, etc., whose nouns are defined as “predicative nouns”. Aiming at further developing the understanding of the syntax-semantics of Brazilian Portuguese (BP) complex sentences, this paper focuses on the understudied phenomenon of phrases such as those underlined in (1):

- (1) a. *Ana (...) declara paixão por futebol*²
(Ana declares [her] passion for football)
b. *Ontem levou, no Rio, os dois filhos a uma visita de rotina ao pediatra.*
(Yesterday [he/she] took, in Rio, the two children to a routine visit to the pediatrician)

In sentence (1a), both the verb *declarar* (“declare”) and the noun *paixão* (“passion”) are autonomous predicates, i.e., they present full-fledged meaning. The former describes a communication, while the latter expresses a feeling. Consequently, their meanings can be separated into two different complete sentences: (i) Ana has a passion for football, and (ii) Ana said that Ana has a passion for football. Similarly, in example (1b), the verb *levar* (“take”) and the predicative noun *visita* (“visit”) represent different events. First, the omitted agent took the children to the visit and, after they arrived, the event of visiting started. In other words, this study focuses on constructions in which a content verb and a predicative noun in object position share a single argument, which can be either the subject or object of the main clause.

In the literature, these constructions have been treated as support verb constructions (Meyers; Reeves; Macleod, 2004) or as control into nominals (Albas-Salas, 2016; Mourounas, 2019), a category which includes support/light verb constructions (including their aspectual variants) in addition to content verb constructions.

Sources such as Meyers, Reeves, and Macleod (2004, p. 97) understand non-light coreferential verbs as (equi) support verbs. However, “support” is typically understood to denote light (i.e., semantically bleached) verbs; “support” and “light” are thus mostly used interchangeably. Similarly, the relation between a light verb and its nominal has been called “control” (Albas-Salas, 2006; Bruening, 2016), as has the relation between aspectual verbs and nominals (Albas-salas, 2006; Mourounas, 2019). Still, sentences like (1) are not comparable to light verb constructions or their aspectual variants. Firstly, because light verbs are not distributional, i.e., they do not select their own subjects (Gross, 1998). Also, light/support verb constructions are often considered to be multi-word expressions or collocations, since verb choice in these constructions often seems arbitrary (e.g. Mel’čuk, 1998; Sag et al., 2002;

¹ No distinction between these terms is made in this paper. “Support” will be used when discussing references which adopt this term and likewise for “light”. The study itself adopts the term “light”, since it is the most widely known.

² All examples in italics provided in this paper are from the Bosque newspaper corpus (AFONSO *et al.*, 2002), except where noted.

Petruck; Ellsworth, 2016; Savary et al., 2017). In contrast, the underlined verbs in (1) denote fully interpretable events and are the nucleus of a compositional phrase.

In turn, aspectual verbs like “begin”, “continue”, and “finish” do not represent predicates *per se*, but rather points in the duration of the noun or clause to which they are syntactically attached. Since they are neither distributional nor entirely compositional (Fotopoulou; Laporte; Nakamura, 2021, p. 8), authors such as Gross (1998) consider them to be support verb variants.

Two general syntactic-semantic accounts of coreference between arguments of both light and content verbs and predicative nouns are known, namely by Albas Salas (2006) and Mourounas (2019). However, a few gaps are observed when it comes to content verbs. The next paragraphs discuss these two proposals and the unaddressed phenomena.

Albas Salas’s (2006) analysis of Spanish, Catalan, and Italian argued for adopting nominal voluntariness and action as the relevant semantic factors for control into nominals:

In fact, the class of verbs that shows obligatory control into nominals (the COMENÇAR- and light FER-types) is only compatible with nouns designating voluntary actions (or potentially voluntary actions, as in the case of Italian *fare una caduta* ‘fall’). By contrast, those verbs that do not show obligatory control into nominals (the DESCRUIRE and PROMETRE class) do not select actional nouns (Alba-Salas, 2006, p. 36).

This generalization can be further discussed when it comes to the phrases under study in BP. Consider sentence (2):

- (2) *A patinadora (...), banida das competições por ter planejado atentado contra a rival (...), será o destaque de setembro da revista (...)*

(The skater, banned from the competitions due to having planned [an] attack against [her] rival, will be the September highlight of the magazine)

In (2), *atentado* (“attack”) is an actional noun and *planejar* (“plan”) is restricted to voluntary activities – it is not normally possible to plan for something to happen accidentally. Still, (2) can be altered into making the attacker and the planner different people, as shown in (3), even if this is a marked usage of this verb phrase.

- (3) O treinador planejou o ataque da patinadora contra a rival
(The coach planned the skater’s attack against her rival)

Paixão in (1a) is non-actional, but that sentence – at least as it is – does not allow for a non-coreferential reading. While this is not ruled out by Albas Salas’s proposal, it remains unexplained why “talk” and “describe” present non-obligatory coreference (p. 12 and 27) – which is also true of its BP counterparts *falar* and *descrever*, as shown in (4) – but *declarar* (“declare”) in (1a) presents a different behavior despite having similar semantics.

- (4) a. Ana (falou sobre + descreveu) a paixão pelo futebol
(Ana (talked about + described) the passion for football)
≠ b. Ana (falou sobre + descreveu) a paixão dela mesma pelo futebol
(Ana (talked about + described) her own passion for football)

In a different passage, Albas Salas equated light verb constructions to obligatory control into nominals and “heavy” (i.e., content) verb constructions to non-obligatory control (p. 16). It is true that, as in sentences (2–3), the attempt to change (1a) into non-coreferential leads to interpretable sentences. Sentence (5) means that Maria, for whatever reason, will not declare her passion herself, and Ana declared it for her, i.e., it has a benefactive interpretation.

- (5) Ana declarou a paixão de Maria pelo futebol
(Ana declared Maria’s passion for football)

Still, the only interpretation of (1a) if unchanged is that *declarar* (“declare”) shares its subject with the predicative noun *paixão* (“passion”). In fact, even light verb constructions can accept these marginal non-coreferential sentences, as noted not only by Albas Salas (p. 24, footnote), but also by recent work on BP support verbs (Rassi, 2023, p. 137). Sentence (6a) provides a light verb construction and (6b) is its non-coreferential version. The latter means that Ana’s passion for football is similar to or as strong as Maria’s.

- (6) a. Ana tem paixão por futebol
(Ana has [a] passion for football)
b. Ana tem a paixão de Maria por futebol
(Ana has Maria’s passion for football)

If *ter paixão* (to have passion) in (6a) counts as an obligatory coreference phrase despite (6b), the acceptability of sentence (5) cannot be used to argue that *declarar paixão* (“declare passion”) is a non-obligatory coreference phrase, and the same is true of the relation between (3) and (2). Instead, this paper will propose that (3), (5), and (6b) are coerced sentences whose unmarked counterparts (2), (1a), and (6a) are always coreferential. This distinction is not only intuitive but also supported by corpus observation since the coerced versions are extremely infrequent.

As an alternative proposal to the syntax-semantics of control into nominals, Mourounas (2019) suggests that obligatory or non-obligatory control into nominals can be explained by whether their complement clause counterparts show “exhaustive control” (EC) or “partial control” (PC). What sets them apart is the fact that PC allows for a collective reading of the subordinate clause even with a singular subject. In English, this can be tested by checking whether a verb with a singular subject takes a subordinate clause with “gather” or the modifier “together”.

- (7) a. *John gathered at 6
b. John wanted to gather at 6
c. John went to the cinema (*together)
d. John wanted to go to the cinema together
(Examples by Mourounas, 2019, p. 27)

Additionally, only PC predicates can have overt subjects in their complement clauses:

- (8) a. Mary hoped that Bill would leave the city
b. Mary hoped for Bill to leave the city
(Examples by Mourounas, 2019, p. 32)

Finally, PC predicates accept tense mismatch between the main verb and the subordinate verb, whereas EC predicates do not. So, in (9), “demand” is a PC verb (9a), since the demanding event can happen yesterday whereas the complement clause, “go to the gym”, takes place tomorrow; “manage” is an EC verb (9b), because “manage” and “go to the gym” cannot happen at different moments.

- (9) a. Yesterday, John demanded to go to the gym tomorrow
 b. *Yesterday, John managed to go to the gym tomorrow
 (Examples by Mourounas, 2019, p. 30)

It turns out, according to Mourounas (2019, p. 34-44), that EC verbs show obligatory control when governing nominals, whereas PC verbs display non-obligatory control under these circumstances. Compare the meanings of (10) and (11).

- (10) a. John (wanted + demanded + planned + hated + disliked) the inspection of the factory
 b. John (wanted + demanded + planned + hated + disliked) to inspect the factory
 (Sentences adapted from Mourounas, 2019, p. 34)

In (10), coreference is not obligatory when the verb selects a nominal (10a), but it is so when selecting an infinitive verb (10b). The matrix verbs of these sentences are PC verbs. This contrasts with the behavior of EC verbs:

- (11) a. John (began + managed) the inspection of the factory
 b. John (began + managed) to inspect the factory
 (Sentences adapted from Mourounas, 2019, p. 35)

The pair of sentences in (11) is paraphrastic; in both cases, it was necessarily John who inspected the factory.

These properties are argued to lead to a semantic clustering of control predicates: adapting other proposals, Mourounas (2019, p. 28–29) suggests that EC (understood as obligatory control into nominals) correlates with “implicative”, “aspectual”, “modal”, and “non-implicative” predicates, whereas PC (non-obligatory control into nominals) correlates with “factive”, “propositional”, “desiderative”, and “interrogative” predicates.

However, not all forms of coreference between arguments of verbs and nouns have a complement clause counterpart which can be assessed for EC or PC. Take, for instance, sentence (12):

- (12) *Os bancos devem responder à proposta do grupo até o dia 12*
 (The banks must respond to the group's proposal by the 12th)

The natural interpretation is that the group proposed something to the bank, i.e., the bank is assigned the recipient role by the predicative noun. However, there is no such possibility as an equivalent complement clause, be it an infinitive with the verbal correspondent of *proposta* (“proposal”), i.e., *propor* (“to propose”) (13a), a finite clause with a light verb construction (13b), or an infinitive clause with a light verb construction (13c):

- (13) a. *Os bancos devem responder a propor até o dia 12
 (*The banks must respond to proposing by the 12th)
 b. ≠ Os bancos devem responder que o grupo fez uma proposta até o dia 12
 (≠ The banks must respond that the group made a proposal by the 12th)
 c. * Os bancos devem responder fazer uma proposta até o dia 12
 (*The banks must respond to make a proposal by the 12th)

Additionally, the predicative noun does play a role, whose importance varies, in how coreference is interpreted. This aspect is not accounted for in a verb-centric classification of these phrases. Constructions with *recusar* (“decline”/ “deny”) are a noteworthy case illustrating how the noun changes semantic role attribution.

- (14) a. *Apesar de evitar dar um não definitivo, Marise deixou claro que deve recusar o convite de Brizola.*
 (Although [she] avoided saying a definitive “no”, Marise made it clear that [she] shall decline Brizola’s invitation)
 b. *A ONU admitiu que os sérvios não haviam cumprido as exigências, mas recusou autorização para bombardeio.*
 (The UN admitted that the Serbs had not complied with the requirements, but denied authorization for bombing)

In (14a), the subject of *recusar* is the recipient of *convite* (“invitation”); in the second, it is the agent of *autorização* (“authorization”). This is further complicated by the fact that both nouns assign both semantic roles and have strikingly similar argument structures³ (double object, one referring to a human being and the other referring to a proposition). These characteristics extend to their verbal counterparts (i.e., *convidar* “invite” and *autorizar* “authorize”). This paper will propose that alternations of this kind can be correlated to acceptance of paraphrases involving different light verb constructions. The noun in (14a) corresponds to a paraphrase with *receber* (“receive”), as in (15a), and (14b) corresponds to the paraphrase with *dar* (“give”) in (15b):

- (15) a. *Marise recusou o convite que recebeu de Brizola*
 (Marise declined the invitation that [she] received from Brizola)
 b. *A ONU recusou dar autorização para bombardeio*
 (The UN refused to give authorization for bombing)

Finally, in BP, the absence of any determiner with certain predicative nouns seems to require a coreferential interpretation (16a–b) which becomes more fluid when an article occurs before the noun (16c). In other words, for certain verb plus noun combinations, coreference seems to be more associated with bare nouns.

³ The prepositions introducing the clausal complements are nonetheless slightly different. *Convite* accepts both *a* and *para*, whereas *autorização* accepts only *para*.

- (16) a. *O Banco apresentou em 1993 um nível de alavancagem alto, demonstrando maior agressividade nas suas operações*
(The Bank presented in 1993 a high level of leverage, showing higher aggressiveness in its operations)
- b. **O Banco demonstrou agressividade dos funcionários*
(*The Bank showed aggressiveness of the workers)
- c. *O Banco demonstrou a agressividade dos funcionários*
(The Bank showed the aggressiveness of the workers)

These empirical observations indicate the need to understand the data on this type of coreference through a wider lens than that of obligatory versus non-obligatory control, or even exhaustive versus partial control, associated with either verbs or nouns alone. This approach should account for how meaning changes under coreferential and non-coreferential uses of these phrases and identify the formal properties related to both readings.

This pilot study departs from previous proposals by adopting a corpus-based approach to argument-sharing content verb plus noun constructions in BP. It draws observations from newspaper corpus examples and pinpoints regularities for further exploration. The objective is to test the utility of the properties pointed out throughout this introduction for classifying these verb phrases and to describe how the nuances of coreference can be related to both verb and nominal semantics.

This introduction presented the singularities of argument-sharing between content verbs and predicative nouns, differentiating them from other types of similar phenomena and showing how current proposals of classification do not fully account for some of their properties. A corpus-based pilot study of this phenomenon was thus proposed. Section 2 discusses the corpus, the theoretical and methodological framework, and the properties selected for analysis. The observed examples and their properties are discussed and systematized in Section 3. Section 4 concludes the findings and proposes further studies on argument-sharing between content verbs and nouns in BP.

2 Method

Since the pilot study presented here stems from an effort to provide an annotation of nouns and their arguments for Natural Language Processing (NLP), the linguistic data were drawn from the Brazilian partition of *Bosque* version 7.5, a subset corpus of the *Floresta Sintá(c)tica* treebank (Afonso et al., 2002)⁴. The Brazilian segment of *Bosque* consists of articles published in the Brazilian newspaper *Folha de São Paulo* in 1994. It has 31,344 tokens, accounting for 8,687 types, distributed into 4,213 sentences.

While this corpus is admittedly small, only partially representative of BP, and not fully adequate for comprehensive corpus research, it was observed to contain data which might contribute to the current understanding of syntactic and semantic constraints of coreference between arguments of verbs and nouns. Considering this potential, the analysis focused on

⁴ Available at <<http://www.linguateca.pt/Floresta/>>, accessed on 01/11/2024.

the 75 verb phrases (VP) which were found to involve coreference of subjects or objects of content verbs with arguments of nouns regardless of whether coreference relied on context or not.

The analysis of corpus sentences and classification proposed in this paper were based on Lexicon-Grammar (Gross, 1975, 1981) as a theoretical and methodological framework. This approach proposes that classification of linguistic items should be founded on a careful registration of their properties on tables and a systematic comparison of their similarities and differences. This paper tests the acceptability of some properties hypothesized to be relevant for the identified VPs to separate them into classes. However, for conciseness, it provides no more than an abridged version of the tables typically employed in Lexicon-Grammar (Table 1). Following the current practices in this framework, the constructions that are acceptable for each lexical item were identified by resorting both to corpora and introspection in a complementary manner: while a given construction emerging in a corpus indicates its acceptability, its absence might just correspond to a gap in the corpus to be filled by introspection (Laporte, 2008). Lexicon-Grammar was employed in recent studies of Brazilian Portuguese, particularly for the description of support verb constructions. These provided a basis for the analysis proposed in this paper. Rodrigues and Picoli (2019, p. 890-895) should be consulted for an overview of these studies.

The semantic and syntactic properties hypothesized to be relevant for coreference – based both on the literature and preliminary observations – were tested for each of the VPs. These were: (1) the different types of semantic coercion which apply when a usually coreferential VP is forced into a non-coreferential reading; (2) the possibilities of substituting the noun with a subordinate clause containing this noun in a light verb construction; (3) how determiners affect coreference; and (4) the possibility of situating the verb and the noun at different moments in time. The following subsections discuss these properties in detail.

2.1 Acceptability of Non-Coreferential Complements and Coercion

While coreference applies for most instances of a given VP, occasional non-coreferential occurrences can be identified. Such infrequent sentences are interpretable only through semantic coercion, i.e., they have an extra meaning component not available in the original, uncoerced sentence.

Such changes in meaning are observed when a non-coreferential argument (either a PP or a possessive pronoun) is added to the original sentence for most of the data on content verbs plus nouns. These changes in meaning are not uniform across all VPs: three types of coercion were observed to apply to this data, in addition to VPs accepting none of them. The first and most recurrent type involves a representative and his or her representee. Consider (17):

- (17) *Os metalúrgicos (...) preparam greve*
(The metalworkers are preparing [for] strike action)

Sentence (17) conveys no other possible interpretation than that the metalworkers are preparing their own strike. However, the agents of *preparar* (“prepare”) and *greve* (“strike action”) might be attributed to different people; it just entails that the former is preparing

the strike for (or in benefit of) the latter. So, (18a) entails (18b). This is named “benefactive coercion” in this paper.

- (18) a. Lula prepara a greve dos metalúrgicos
(Lula prepares the metalworkers’ strike action)
b. => Lula prepara a greve dos metalúrgicos (para + por) eles
(Lula prepares the metalworkers’ strike action for them)

A second type of coercion involves a discrepancy between the real world and a reported scenario. It applies to certain predicates denoting transfer of information and, when coreference is broken, entails that such information is incorrect. For example, sentence (19) is neutral regarding whether the fisherpersons committed a crime or not.

- (19) *Os pescadores (...) serão acusados de crime contra fauna*
(The fisherpersons will be accused of crime against the fauna)

If *empresa* (company) is added as an agent for *crime* (20a), the sentence now entails that the fisherpersons did not commit the crime, that the company did, and that they are being accused unfairly (20b). This will be referred to as the “counterfactual coercion”.

- (20) a. *Os pescadores serão acusados do crime da empresa contra a fauna*
(The fisherpersons will be accused of the company’s crime against the fauna)
b. => *Quem cometeu o crime foi a empresa, não os pescadores*
(It was the company that committed the crime, not the fisherpersons)

A third type of coercion implies that the event denoted by the noun affects the main clause’s subject psychologically.

- (21) *Eddie (...) não resiste à tentação da atriz*
(Eddie doesn’t resist the actress’ temptation)

Inserting a different experiencer complement for *tentação* (“temptation”) in sentence (22a) suggests that the temptation is not directed at Eddie but affects him somehow (22b). In this paper this is referred to as “affective coercion”.

- (22) a. *Eddie não resiste à tentação da atriz contra o Paulo*
(Eddie does not resist the temptation of the actress against Paulo)
b. => *A tentação da atriz contra o Paulo afeta o Eddie psicologicamente*
(The temptation of the actress against Paulo affects Eddie psychologically)

Finally, there are verb plus noun combinations that allow an uncoerced interpretation of lack of coreference. This applies to (23); as it is phrased, it suggests that football (metonymically understood as people somehow involved in football) should present all the wrongdoings of football itself, just like the parliamentary inquiry commission presented its own wrongdoings.

- (23) *O futebol precisa seguir o exemplo da CPI do orçamento e apresentar todos os podres*
 (Football needs to follow the example of the budget parliamentary inquiry commission and present all the wrongdoings)

Changing background information and inserting an argument for *podres* (literally “rotten things”, interpreted here as “wrongdoings”) disrupts the coreferential interpretation.

- (24) *Os jornalistas precisam apresentar todos os podres da Ana*
 (The journalists must present all of Ana’s wrongdoings)

However, this syntactic similarity with the previous cases hides an important semantic distinction: from (24) it does not follow that the journalists are representing Ana in any way, or that Ana is the one who actually has flaws while people attribute them to journalists, or that Ana’s flaws affect the journalists psychologically. Thus, the interpretation of this type of sentence does not follow any of the pre-established patterns commonly observed for lack of coreference. This is a true non-obligatory coreference construction, in the sense that lack of coreference leads to an unmarked sentence. Coreference in these cases is possibly due to contextual factors and might be unrelated to verb or noun semantics.

2.2 Paraphrases with Light Verb Constructions

Gross (1981, 1998) proposed that predicative nouns are characterized by their relations with support verb constructions, even when they occur in other environments. In content verb constructions, the predicative NP (25d) can alternate between a complement clause with a light verb construction (25a) and an NP containing a relative clause with a light verb construction (25b-c).

- (25) a. *Luc apprécie que ce texte ait une certaine élégance*
 (Luc appreciates that this text has a certain elegance)
 b. *Luc apprécie l’élégance que ce texte a*
 (Luc appreciates the elegance that this text has)
 c. *Luc apprécie l’élégance de ce texte*
 (Luc appreciates the elegance of this text)
 (Examples adapted from Gross, 1998, p. 18)

Not all VPs correspond to all these paraphrases. This varies with the verb’s selection properties, which can be exemplified with the VPs *acatar a recomendação* (“accept the recommendation”) and *negar envolvimento* (“deny involvement”). In the former, the noun corresponds to its own version with a relative clause containing a light verb construction (26a-b), but not a complement clause (26c-d).

- (26) a. (...) *Mário (...) acatou as recomendações de assessores do presidente eleito (...)*
 (Mário accepted the recommendations of the president-elect’s advisors)
 b. *Mário acatou as recomendações que os assessores do presidente eleito fizeram*

- (Mário accepted the recommendations that the president-elect's advisors made)
- c. *Mário acatou que os assessores do presidente eleito tenham feito recomendações
(*Mário accepted that the president-elect's advisors have made recommendations)
- d. *Mário acatou os assessores do presidente eleito fazerem recomendações
(*Mário accepted the president-elect's advisors to make recommendations)

Negar envolvimento works the other way around: it accepts the substitution of the noun with a complement clause (27c-d), but not with the noun plus a relative clause (27b). While (27b) is acceptable, it entails the truth of the relative clause and therefore conveys more information than (27a), which makes no judgement on whether the group and the Iranian government are involved or not.

- (27) a. *O grupo e o governo iraniano negaram envolvimento*
(The group and the Iranian government denied involvement)
- b. ≠ *O grupo e o governo iraniano negaram o envolvimento que tinham*
(≠ The group and the Iranian government denied the involvement that they had)
- c. *O grupo e o governo iraniano negaram que tinham envolvimento*
(The group and the Iranian government denied that they had involvement)
- d. *O grupo e o governo iraniano negaram ter envolvimento*
(The group and the Iranian government denied having involvement)

The possibility of substituting nouns with (finite and non-finite⁵) light verb complement clauses or relative clauses was hypothesized to be correlated with different types of coreference. These were marked as either paraphrastic or non-paraphrastic for each of the VPs.

Six light verbs were found to account for possible paraphrases of these VPs: *fazer* (“to make/do”), *ter* (“to have”), *receber* (“to receive”), *sofrer* (“to suffer”), *cometer* (“to commit”), and *dar* (“to give”). Each of these verbs seems to have some correlation with semantic features. Butt (2010) suggested that cross-linguistically light verb choice cuts between agentive (“to make/do”, “to give”) and non-agentive (“to have”) subject semantic role. For BP, Barros’s (2014) work suggests that most nouns combining with *fazer* are actional – although certain nouns, such as names of exams and medical procedures, may present a patient subject (p. 160–161). Most of Rassi’s (2023, p. 157–158) classes of constructions with light verb *dar* also present agentive subjects. Other verbs (to receive, to suffer) have been described (G. Gross, 1989) as passive alternations of “to do/make” and “to give”, placing the patient of the noun as their subject. Calcia (2022) provides a further examination of these verbs in BP; *sofrer* tends to occur with nouns having a negative connotation. *Ter* has also been included as a converse construction for action nouns. Based on Santos’ (2015) list of nouns occurring with *ter*, non-actional nouns

⁵ An additional property was initially included based on the literature distinguishing partial and exhaustive control (Mourounas, 2019): the possibility of an overt subject in infinitive clauses, as briefly discussed in the Introduction. This was quickly shown to lead to sparse results, as very few phrases presented such property, making it difficult to discuss in depth for this pilot study’s purposes, and was thus dropped at a later stage. It is not fully ruled out that this property might be useful for a more comprehensive account of this phenomenon.

combining with this verb seem to usually include names of diseases, psychological states, and qualities. Finally, “to commit” is an agentive verb with negative connotations (Gross, 1998, p. 39–40). Light verb acceptability is thus a somewhat reliable formal indicator of certain semantic properties, which makes it of importance for the classifications of nouns if their irregularities are considered.

2.3 Determiners

During preliminary analyses, some VPs seemed to be more related to a coreferential interpretation when the noun lacked a determiner. This is exemplified in (28).

- (28) a. (...) *Caetano (...) manifestou desejo de assistir ao show (...)*
(Caetano expressed [his] desire to watch the show)
b. *Caetano manifestou (o + seu) desejo de assistir ao show*
(Caetano expressed (the + his) desire to watch the show)

In the original sentence (28a), a coreferential interpretation of *manifestar desejo* (“manifest desire”) is obtained with no determiner before the noun. The same interpretation is possible with a definite article and/or coreferential possessive pronoun before the noun (28b). Coreference in cases like (28b) can be neutralized if the noun’s experiencer argument is filled with a PP (29a) or possessive pronoun (29b):

- (29) a. *Caetano manifestou o desejo de Gil de assistir ao show*
(Caetano manifested the desire of Gil to watch the show)
b. *Caetano manifestou o meu desejo de assistir ao show*
(Caetano manifested my desire to watch the show)

However, VPs like (29a) do not easily accept determiner omission under this non-coreferential interpretation (30):

- (30) ?*Caetano manifestou desejo de Gil de assistir ao show*
(Caetano manifested desire of Gil to watch the show)

A potential argument against including this property is that relying on introspection for subtle differences may introduce biased judgments. Objectivity can be enhanced by quantitatively comparing the frequency of VPs with and without determiners in larger corpora, considering both coreferential and non-coreferential readings. For example, in Corpus Brasileiro (SARDINHA, 2009)⁶, out of 373 instances of *manifestar o desejo* (“manifest the desire”), only two involve a clear experiencer NP complement for *desejo* (“desire”).

- (31) a. (...) *tive a singular oportunidade de manifestar o desejo do povo brasileiro da construção de um futuro melhor para todas as nações (...)*

⁶ Lemma frequency consulted with Linguateca’s AC/DC tool, available at <<https://www.linguateca.pt/>>. Last access: 06/14/2023.

(I had the singular opportunity of manifesting the desire of the Brazilian people for the construction of a better future for all nations)

- b. *Certamente, a iniciativa popular é a marca principal desse projeto, que veio para esta Casa, por meio da coleta de milhares de assinaturas, manifestando o desejo da população brasileira de ter melhores condições habitacionais.*

(Certainly the popular initiative is the main mark of this project, which came to this House through the collection of thousands of signatures, manifesting the desire of the Brazilian population for better housing conditions)

(Corpus Brasileiro)

The version without the article, *manifestar desejo* (“manifest desire”), has no instances with a clear experiencer NP complement in Corpus Brasileiro, although this VP occurs 140 times.

2.4 Tense Mismatch

Mourounas (2019, p. 38-39) considered tense mismatch to be a property of partial control predicates. According to this author, the sentence in (32a) cannot accept tense mismatch because it presents exhaustive control, whereas (32b) presents partial control.

- (32) a. *Yesterday, John (began + resumed + managed + tried) the inspection of the factory tomorrow
b. Yesterday, John (wanted + demanded + planned) the inspection of the factory tomorrow
(Examples adapted from Mourounas, 2019, p. 37-38)

Tense mismatch can be either future- or past-oriented. The former enables time modification of the nominal to a subsequent moment in relation to that of the verb, as in (32b). Past-oriented tense mismatch works the other way around, accepting only nominals interpreted as preceding the verb’s event (33a):

- (33) a. Today, John (hated + regretted + disliked) the inspection of the factory yesterday
b. *Today, John (hated + regretted + disliked) the inspection of the factory tomorrow
(Examples adapted from Mourounas, 2019, p. 39)

3 Classification

It has long been said that rarely two predicates present the same properties (e.g., Gross, 1975, p. 214). This is also true of the VPs under investigation in this paper: most present unique patterns in one way or another and can only be grouped together based on similarity, not

identity. The following subsections try to group the VPs under analysis into classes that seem syntactically and semantically coherent regarding the properties described in section 2. The classification criteria vary from one class to another, since some properties crucial for characterizing some VPs are irrelevant for others.

The names of the classes are phrased to provide clues on the semantics of the VPs as observed in the corpus and should not be taken as definitive. What truly characterizes the VPs are the formal properties discussed in each subsection, and the underlying notional semantics of these classes remains open for debate as further studies are conducted.

3.1 Intention and Fulfillment Verbs Plus Action Nouns

A major group of predicates was found to have similar behaviors regarding the properties discussed in section 2. These are noted to lead to the benefactive coercion when a non-coreferential interpretation is forced upon the sentence, e.g., (34b) entails (34c).

- (34) a. *Zanini diz que está preparando a fundação de um novo partido (...)*
(Zanini says that he is preparing the foundation of a new party)
b. ?*Zanini está preparando a fundação de um partido pela Ana*
(Zanini is preparing the foundation of a party by Ana)
c. = ?*Zanini está preparando a fundação de um partido pela Ana para ela*
(Zanini is preparing the foundation of a party by Ana for her)

They also accept tense mismatch between verb and noun:

- (35) *Zanini está preparando hoje a fundação de um partido amanhã*
(Zanini is preparing today the foundation of a party tomorrow)

The addition of a subordinate clause containing light verb *fazer* (“do/make”)⁷ is mostly possible, even when it results in an interpretation not inherent in the noun alone, as in (36).

- (36) *Zanini está preparando a fundação do partido que ele vai fazer*
(Zanini is preparing the foundation of the party that he will make)

Oferecer cursos (“offer courses”) has almost the same characteristics, except that instead of *fazer* it forms paraphrases with light verb *dar* (“to give”), which was shown to be a generally agentive light verb (Rassi, 2023, p. 157–158)⁸.

⁷ Some members of this class also accept *ter* (have) in addition to *fazer*, but this seems to be irrelevant for the classification.

⁸ This is not to say that *fazer* and *dar* are mere collocational variations of one another: *dar um curso* translates to “teach a course”, whereas *fazer um curso* corresponds to “take a course”. They denote different forms of agentivity, and *oferecer cursos* can only correspond to the former.

- (37) a. *Nos Estados Unidos, universidades e algumas empresas chegam a oferecer cursos de boas maneiras nos negócios*
 (In the United States, universities and some companies even offer courses on good manners in business)
- b. *As universidades oferecem os cursos que vão dar no próximo semestre*
 (The universities offer the courses they will give[=teach] next semester)

Nineteen VPs demonstrate coherence with these properties: *acertar dívida* (“pay the debt”), *buscar independência* (“search for independence”), *concordar com a transferência* (“agree with the transference”), *conclamar a uma jihad* (“call for a jihad”), *cumprir a promessa* (“fulfill the promise”), *encomendar o desenvolvimento* (“order the development”), *investir em inaugurações* (“invest in inaugurations”), *negociar compra* (“negotiate [a] purchase”), *negociar um tratado* (“negotiate an agreement”), *oferecer cursos* (“offer courses”), *optar pela locação* (“opt for the rental”), *planejar atentado* (“plan [an] attack”), *preparar a fundação* (“prepare the foundation”), *preparar greve* (“prepare [for] strike action”), *preparar o discurso* (“prepare the discourse”), *prometer um anúncio* (“promise an announcement”), *realizar essas metas* (“meet these goals”), *tentar a libertação* (“try the liberation”), *tentar contato* (“try contact”).

These VPs diverge in whether their tense mismatch is past- or future-oriented and whether their relative clause version serves as a paraphrase of the noun alone. Past-oriented tense mismatch is associated to a paraphrastic reading of the relative clause, whereas future-oriented tense mismatch corresponds to a non-paraphrastic reading of noun plus relative clause. This distinction divides this major class into the two classes discussed in the following subsections.

3.1.1 Intention Verbs Plus Action Nouns

This class includes VPs of the preceding discussion that accept future-oriented tense mismatch, as shown in (35), and reject past-oriented tense mismatch (38):

- (38) **Hoje Zanini está preparando a fundação de um partido ontem*
 (*Today Zanini is preparing the foundation of a party yesterday)

In addition, relative clauses with *fazer* (do/make) imply either a factive reading or that an event is believed by the speaker to happen in the future, which is not true of the sentence with the noun alone. For example, (34a) does not imply (36), since Zanini might end up not founding the party at all. Consequently, the inclusion of an adversative clause is pragmatically acceptable for (34a), but not so much for (36), as shown in (39):

- (39) a. *Zanini está preparando a fundação de um partido, mas ele não vai conseguir fundar esse partido*
 (Zanini is preparing the foundation of a party, but he won't be able to found this party)
- b. *#Zanini está preparando a fundação de um partido que ele vai fazer, mas ele não vai conseguir fundar esse partido*
 (#Zanini is preparing the foundation of a party that he will make[=establish], but he won't be able to found this party)

Sixteen VPs exhibit these two properties: *buscar independência* (“search for independence”), *concordar com a transferência* (“agree with the transference”), *conclamar a uma jihad* (“call for a jihad”), *encomendar o desenvolvimento* (“order the development”), *investir em inaugurações* (“invest in inaugurations”), *negociar compra* (“negotiate [a] purchase”), *negociar um tratado* (“negotiate an agreement”), *oferecer cursos* (“offer courses”), *optar pela locação* (“opt for the rental”), *planejar atentado* (“plan [an] attack”), *preparar a fundação* (“prepare the foundation”), *preparar greve* (“prepare [for] strike action”), *preparar o discurso* (“prepare the discourse”), *prometer um anúncio* (“promise an announcement”), *tentar a libertação* (“try the liberation”), and *tentar contato* (“try contact”).

3.1.2 Fulfillment Verbs Plus Action Nouns

These are the opposite of what characterized the previous class. They accept past-oriented (40b), but not future-oriented (40c) tense mismatch:

- (40) a. *Agricultor acerta dívida*
(Farmer pays debt)
b. *Hoje agricultor acerta dívida de ontem*
(Today farmer pays yesterday’s debt)
c. **Hoje agricultor acerta dívida de amanhã*
(*Today farmer pays tomorrow’s debt)

Both noun alone (41a) and noun plus relative light verb construction (41b) have factive readings, and are thus paraphrastic; denying that the farmer incurred debt leads to an infelicitous utterance in both cases:

- (41) a. #*O agricultor acertou a dívida dele, mas ele não fez essa dívida*
(#The farmer paid his debt, but he didn’t make[=incur] this debt)
b. #*O agricultor acertou a dívida que fez, mas ele não fez essa dívida*
(#The farmer paid the debt that [he] made[=incurred], but he didn’t make[=incur] this debt)

Only three VPs presented these properties: *acertar dívida* (“pay debt”), *cumprir a promessa* (“fulfill the promise”), and *realizar essas metas* (“meet these goals”).

3.2 Credit Verbs Plus Dysphoric Action Nouns

These VPs were found to have properties in common regarding type of coercion and paraphrases with subordinate light verb constructions. Firstly, they were the only VPs to accept the counterfactual coercion:

- (42) *Ele é acusado de vandalismo.*
(He is accused of vandalism)

Trying to insert the agent of *vandalismo* (“vandalism”) into the NP in (42) leads to the interpretation that “he” did not commit this vandalism, but this argument within NP did, as shown in (43).

- (43) a. Ele é acusado do vandalismo da Maria
(He is accused of Maria’s vandalism)
b. => Quem cometeu esse vandalismo foi a Maria, não ele
(It was Maria who committed this vandalism, not him)

Also, these nouns usually accept being paraphrased with a light verb clause containing *cometer* (“to commit”), although the type of subordinate clause varies among predicates.

- (44) Ele é acusado de ter cometido vandalismo
(He is accused of having committed vandalism)

Finally, past-oriented tense mismatch is applicable to all of them:

- (45) Amanhã Maria vai ser acusada do vandalismo de ontem
(Tomorrow Maria will be accused of yesterday’s vandalism)

Future-oriented tense mismatch is not available for any of these predicates:

- (46) *Ontem Maria foi acusada do vandalismo de amanhã
(Yesterday Maria was accused of tomorrow’s vandalism)

Only six VPs were found to present this behavior: *acusar de* (“accuse of”) *crime*, *estupro* (“rape”), *morte* (“death”), and *vandalismo* (“vandalism”), *atribuir o atentado* (“attribute the attack”), and *confessar o crime* (“confess the crime”).

3.3 Demand and Response Verbs Plus Nouns

A certain type of light verb alternation similar to the active-passive voice alternation was named “conversion” by G. Gross (1989). This work inspired analyses of these constructions in BP, such as Calcia’s (2022). The latter author proposes that these constructions can be divided into five different syntactic-semantic classes; the VPs observed herein seem to correspond mostly to Calcia’s *fazer-receber* (“do/make-receive”) and *fazer-sofrer* (“do/make-suffer”) classes. The examples in (47a-b) and (47c-d) illustrate the process of conversion that characterizes these two classes, respectively.

- (47) a. O laboratório fez uma proposta ao governo brasileiro
(The laboratory made a proposal to the Brazilian government)
b. => O governo brasileiro recebeu uma proposta do laboratório
(The Brazilian government received a proposal from the laboratory)
c. O conselho fez uma injustiça ao coronel
(The council made[=perpetrated] an injustice against the colonel)

- d. => O coronel sofreu uma injustiça do conselho
 (The colonel suffered an injustice from the council)
 (Sentences from Calcia, 2022, p. 582)

Eighteen VPs were verified to constitute a major class due to accepting paraphrases with converse constructions with *receber* (“to receive”) and/or *sofrer* (“to suffer”). These were *acatar a recomendação* (“accept the recommendation”), *atender encomenda* (“meet an order/request”), *cumprir as exigências* (“meet the requirements”), *ceder à pressão* (“give in to pressure”), *condenar à morte* (“condemn to death”), *esperar resposta* (“wait for [an] answer”), *merecer atenção* (“deserve attention”), *pedir ajuda* (“ask for help”), *pedir apoio* (“demand support”), *pedir folga* (“ask for time off”), *pedir uma indenização* (“demand a compensation”), *precisar de (a/uma) ajuda* (“need the/a help”), *recusar o convite* (“decline the invitation”), *responder perguntas* (“answer questions”), *responder à proposta* (“answer the proposal”), *resistir à tentação* (“resist the temptation”), *suportar a pressão* (“handle the pressure”), *usar a pressão* (“use the pressure”). The sentences in (48) exemplify the possibility of paraphrases in these VPs.

- (48) a. *O desembargador (...) vai pedir apoio do Exército para subir morros (...)*
 (The judge will ask for the army’s support to ascend [the] hills)
 b. *O desembargador vai pedir para receber apoio do Exército para subir morros*
 (The judge will ask to receive support from the army to ascend [the] hills)
 c. (...) *Marise deixou claro que deve recusar o convite de Brizola*
 (Marise made it clear that she shall decline Brizola’s invitation)
 d. *Marise deixou claro que deve recusar o convite que recebeu de Brizola*
 (Marise made it clear that she shall decline the invitation that she received from Brizola)

These VPs were found to form three groups with common syntactic and semantic properties plus a *hapax legomenon*. Each of the following subsections is dedicated to one of these groups.

3.3.1 Demand Verbs Plus Nouns

Demand constructions stand out from the other two due to accepting being paraphrased with an infinitive clause, as exemplified in (48b). They also accept relative clauses (49) with a factive reading imprinted on the noun, while the original sentence (48a) is vague in relation to whether the event denoted by the noun happened or not:

- (49) *O desembargador pediu o apoio que recebeu do exército*
 (The judge asked for the support that [he] received from the army)

Complementarily, attempts at creating non-coreferential sentences with these VPs lead to the benefactive coercion:

- (50) a. O desembargador pediu apoio do exército à Maria⁹
 (The judge asked for [the] army's support to Maria)
 b. => O desembargador pediu apoio do exército à Maria para ela
 (The judge asked for the army's support to Maria for her)

Only future-oriented tense mismatch is available for these VPs:

- (51) Hoje, o desembargador pediu apoio do exército mês (que vem + *passado)
 (Today, the judge asked for the army's support (next + *last) month)

The VPs matching these criteria amounted to seven: *esperar resposta* ("wait for [an] answer"), *merecer atenção* ("deserve attention"), *pedir ajuda* ("ask for help"), *pedir apoio* ("ask for support"), *pedir folga* ("ask for time off"), *pedir uma indenização* ("demand a compensation"), *precisar de a/uma ajuda* ("need the/a help").

3.3.2 Response Verbs Plus Request Nouns

Response to request VPs were observed to be remarkably similar to demand pairs, since both can be paraphrased with a construction with *receber* ("to receive") and present the benefactive coercion. These differ from the former, firstly, due to not taking infinitive clauses and because both the noun alone and its version with a relative clause present a factive reading. We see in (52) that (48c) cannot correspond to an infinitive clause; although the sentence itself is not completely unacceptable (though hard to contextualize), it does not correspond to the original, since it implies that Marise did not even receive the invitation.

- (52) ?Marise deixou claro que deve recusar receber o convite de Brizola
 (?Marise made clear that [she] shall decline to receive Brizola's invitation)

They further differ from demand plus noun VPs due to presenting only past-oriented tense mismatch:

- (53) Hoje Marise recusou o convite (de ontem + *de amanhã)
 (Today Marise declined (yesterday's + *tomorrow's) invitation)

This behavior was observed in six VPs: *acatar a recomendação* ("comply with the recommendation"), *atender encomenda* ("meet an order/request"), *cumprir as exigências* ("meet the requirements"), *recusar o convite* ("decline the invitation"), *responder à proposta* ("answer the proposal"), *responder perguntas* ("answer questions").

⁹ In this ambiguous sentence, *à Maria* ("to Maria") can be interpreted either as an argument of *pedir* ("request") or *apoio* ("support"). This analysis considers only the latter interpretation.

3.3.3 Reaction Verbs Plus Dysphoric Action Nouns

Four VPs were characterized by accepting paraphrasis with a relative clause containing light verb *sofrer* (“suffer”) in addition to *receber* (“receive”), maintaining a factive reading, as shown in (54b), but not with a complement clause (54c):

- (54) a. *Culpi, no entanto, acha que o time pode suportar a pressão (...)*
(Culpi, however, believes that the team can handle the pressure)
b. *Culpi acha que o time pode suportar a pressão que está sofrendo*
(Culpi believes that the team can handle the pressure that [it] is suffering)
c. **O time pode suportar que está sofrendo pressão*
(The team can handle that [it] is suffering pressure)

These VPs are also coherent in their acceptance of the affective coercion:

- (55) a. *Culpi suportou a pressão do rival contra o time*
(Culpi handled the pressure of the rival against the team)
b. => *A pressão do rival contra o time afeta Culpi psicologicamente*
(The rival’s pressure against the team affects Culpi psychologically)

Finally, tense mismatch leads to sentences with very low acceptability in this class:

- (56) ??*Hoje o time suporta a pressão de (ontem + amanhã)*
(Today the team handles (yesterday’s + tomorrow’s) pressure)

Four combinations were found to have these properties: *ceder à pressão* (“give in to the pressure”), *resistir à tentação* (“resist the temptation”), *suportar a pressão* (“handle the pressure”), and *usar a pressão* (“use the pressure”).

The properties of *condenar à morte* (“condemn to death”) are remarkably similar both to “demand verbs plus nouns” and “reaction verbs plus dysphoric action nouns”. It differs from the latter by also accepting light verb *ter* (“have”) in addition to *sofrer* and an infinitive complement clause paraphrase, which is what draws it closer to the former set:

- (57) a. *Miami condena [pessoa] a 5 mortes*
(Miami condemns person to 5 deaths)
b. *Miami condena pessoa a (sofrer + ter) uma morte (indolor + horrível)*
(Miami condemns person to (suffer + have) a (painless + horrible) death)

At the same time, *condenar à morte* is subject to affective coercion just like the latter class. Sentence (58) can only make sense if the mother is psychologically affected by her son’s death.

- (58) *Miami condena mãe à morte do filho*
(Miami condemns mother to death of [her] son)

3.4 Manifestation Verbs Plus Psychological Nouns

These VPs accept both complement clause (59b-c) and relative clause (59d) paraphrases with *ter* (“have”). The latter sometimes has a factive reading not inherent in the original:

- (59) a. *Não é permitido falar muito contra a monarquia no Parlamento porque supostamente você deve jurar fidelidade à rainha*
(It is not permitted to speak too much against the monarchy in the Parliament because supposedly you must swear loyalty to the queen)
- b. *Você deve jurar que tem fidelidade à rainha*
(You must swear that [you] have loyalty to the queen)
- c. *Você deve jurar ter fidelidade à rainha*
(You must swear to have loyalty to the queen)
- d. *Você deve jurar a fidelidade que tem à rainha*
(You must swear the loyalty that [you] have to the queen)

The coreferential reading seems to be associated with the bare noun, as sentences without coreference do not sound as natural without a determiner. This is the only class in which determiners, as discussed in 2.3., are shown to play a role in coreference.

- (60) a. ??*Você deve jurar fidelidade do seu filho à rainha*
(You must swear loyalty of your son to the queen)
- b. *Você deve jurar a fidelidade do seu filho à rainha*
(You must swear the loyalty of your son to the queen)

These are the only properties that hold these predicates together as a class. Some present the benefactive semantic coercion and some simply do not have a coerced interpretation. *Jurar fidelidade* (“swear loyalty”) falls into the first case, as evidenced by (60a) entailing (61):

- (61) *Você deve jurar a fidelidade do seu filho à rainha por ele*
(You must swear the loyalty of your son to the queen for him)

Demonstrar agressividade (“show aggressiveness”), in turn, is only non-coreferential in a sense involving someone or a situation providing evidence that someone else is aggressive without necessarily acting as its representative:

- (62) a. *O Banco apresentou em 1993 um nível de alavancagem alto, demonstrando maior agressividade nas suas operações*
(The bank presented in 1993 a high level of leverage, showing higher aggressiveness in its operations)
- b. *O palestrante demonstrou a agressividade do banco*
(The speaker showed the bank’s aggressiveness)

These VPs were not considered to be examples of true non-obligatory coreference because, when the article is missing, coreference seems to be obligatory (63), and lack of coreference only applies to the version with the article (62b).

- (63) * O palestrante demonstrou agressividade do banco
(The speaker showed bank's aggressiveness)

Tense mismatch under a coreferential reading leads to sentences ranging from clumsy to unacceptable:

- (64) a. Hoje Ana demonstrou paixão (??de ontem + *de amanhã) por futebol
(Today Ana showed (yesterday's + tomorrow's) passion for football)
b. *Hoje o banco demonstrou maior agressividade (ontem + amanhã)
(Today the bank showed higher aggressiveness (yesterday + tomorrow))

The class examined in this subsection includes seven VPs: *declarar paixão* ("declare passion"), *declarar simpatia* ("declare sympathy"), *demonstrar agressividade* ("show aggressiveness"), *manifestar desejo* ("manifest desire"), *manifestar preocupação* ("manifest preoccupation"), *mostrar disposição* ("show disposition"), *jurar fidelidade* ("swear loyalty").

As this subsection concludes the analysis of lexically motivated coreference in the VPs, Table 1 provides a synthesis of the classification presented so far.

Table 1 – Classes of Content Verb Plus Predicative Noun Phrases with a Preference for Coreferential Readings

Class	Coercion	Tense Mismatch	Paraphrastic Light Verb	N = Finite Clause	N = Infinitive Clause	N = Relative Clause	Bare Noun Triggers Coreference
Intention Verbs Plus Action Nouns	Benefactive	Future-oriented	<i>Fazer</i> ("make/do") or <i>fazer</i> and <i>ter</i> ("have") or <i>dar</i> ("give")	Varies	Varies	No	No
Fulfillment Verbs Plus Action Nouns	Benefactive	Past-oriented	<i>Fazer</i> ("make/do") or <i>fazer</i> and <i>ter</i> ("have")	Varies	Varies	Yes	No
Credit Verbs Plus Dysphoric Action Nouns	Counterfactual	Past-oriented	<i>Cometer</i> ("commit")	Varies	Varies	No	No
Demand Verbs Plus Nouns	Benefactive	Future-oriented	<i>Receber</i> ("receive")	Yes	Varies	No	No
Response Verbs Plus Request Nouns	Benefactive	Past-oriented	<i>Receber</i> ("receive")	No	No	Yes	No

Reaction Verbs Plus Dysphoric Action Nouns	Affective	Nearly unacceptable	<i>Sofrer</i> (“suffer”) and <i>receber</i> (“receive”)	No	No	Yes	No
Manifestation Verbs Plus Psychological Nouns	Benefactive or none	Nearly unacceptable	<i>Ter</i> (“have”)	Yes	Yes	Varies	Yes

Source: The author

3.5 True non-obligatory coreference

Lack of coreference did not lead to a coerced interpretation in the same way as for the previous VPs – even accounting for determiner variation – in VPs such as the following:

- (65) (...) *nem os grandes países souberam resolver o problema do desemprego*
 (Not even the big countries were able to solve the problem of unemployment)

If we try to build a sentence in which the problem of unemployment is not a characteristic of the big countries, we get sentence (66). This is arguably more acceptable than the coerced sentences seen so far, suggesting that it is not a coerced sentence (i.e., it is unmarked).

- (66) *Os grandes países resolveram o problema dos países menores*
 (The big countries solved the problem of the smaller countries)

The benefactive coercion might seem applicable to this specific case. However, this is not a necessary interpretation. Consider sentences with *resolver o problema* (“solve the problem”) with non-human subjects:

- (67) *A melhora na economia resolveu o problema dos países menores*
 (The improvement in the economy solved the problem of the smaller countries)

The phrase “the improvement in the economy” lacks agency, which makes sentence (68) odd:

- (68) ??*A melhora na economia resolveu o problema dos países menores para eles*
 (The improvement in the economy solved the problem of the smaller countries for them)

Out of 10,309 occurrences of *resolver o problema* (“solve the problem”) in Corpus Brasileiro, 4,944 are followed by a PP introduced with *de* (of), generally corresponding to nom-

inal complements. Therefore, both intuition and corpora tell us that the benefactive coercion is not applicable here, placing *resolver o problema* as a true non-obligatory coreference VP.

Most of these nouns can also be paraphrased with a relative clause containing a light verb. However, *vender a imagem* (“sell the image”) and *temer cansaço* (“fear tiredness”) differ from the rest because the relative clause will add a factive interpretation.

- (69) a. (...) *um jornal que vende a imagem de um dos mais cosmopolitas do mundo*
(A newspaper that sells the image of [being] one of the most cosmopolitan in the world)
- b. ≠ *O jornal vende a imagem de cosmopolita que ele tem*
(≠ The newspaper sells the image of cosmopolitan that it has)

Other than this, there is very little syntactic homogeneity among these pairs, which – if we are to accept that similar form is related to similar meaning – reinforces the idea that they do not form any type of coherent lexical group.

The 17 VPs matching these characteristics are *admitir a possibilidade* (“admit the possibility”), *afastar-se da pesquisa* (“move away from the research”), *anunciar sua candidatura* (“announce one’s candidacy”), *aproveitar a viagem para* (“take advantage of the trip to”), *atrair apoio* (“draw support”), *descartar a candidatura* (“discard the candidacy”), *encontrar um discurso* (“find a discourse”), *apresentar os podres* (“present the wrongdoings”), *irritar-se com uma pergunta* (“be irritated at a question”), *conformar-se com o papel* (“conform with the role”), *levar a uma visita* (“take to a visit”), *negar envolvimento* (“deny involvement”), *renunciar à candidatura* (“renounce the candidacy”), *resolver o problema* (“solve the problem”), *salvar da ignorância* (“save from ignorance”), *temer cansaço* (“fear tiredness”), and *vender a imagem* (“sell the image”).

3.6 Outliers

As is bound to happen in linguistic studies in general – particularly a pilot study – some of the observations did not seem to match any class. The next paragraphs present a summary discussion of these standalone cases.

Concentrar as críticas (“concentrate criticism”), *assinar o tratado* (“sign the treaty”), and *assinar o acordo* (“sign the deal”) do not accept tense mismatch and cannot be paraphrased with complement clauses. This makes them also somewhat similar to “response verbs plus dysphoric action noun”, although they cannot be categorized as such. These three combine both with light verbs *fazer* (“make/do”) and *ter* (“have”), similarly to certain nouns of the “intention and fulfillment” major class. *Avançar nas negociações* (“advance in the negotiations”) has a similar behavior, except that it cannot be paraphrased with *ter* and only accepts tense mismatch as long as the “advancing” event is included in the “negotiating” event, which prevents it from classification as “intention verb plus action noun”. These factors suggest that these might constitute a class, but no notional semantic coherence can be found for these four examples, two

of which have the same verb (*assinar*) and nouns that are very close in meaning (i.e., *tratado* and *acordo*). Also, *assinar* corresponds to non-distributional verbs (Gross, 1998, p. 15).

Confirmar a vantagem (“confirm the advantage”) almost fits into “manifestation verbs plus psychological nouns”: it accepts subordinate clauses with light verb *ter*. However, it deviates from the typical behavior of this class since the noun here is more acceptable with an article:

(70) *Culpi, no entanto, acha que o time pode suportar a pressão e “confirmar a vantagem” em São Paulo*¹⁰

(Culpi, however, thinks that the team can handle the pressure and “confirm the advantage” in São Paulo)

The version without the article is not uninterpretable or unattested, but it is overwhelmingly less common, as can be confirmed by a simple corpus query:

(71) ?O time pode confirmar vantagem
(?The team may confirm advantage)

Mudar de vida (“change [one’s own] life”) (72a) and *recusar uma postura* (“refuse an attitude”) (72c) lie in the scarcely populated end of the coreference spectrum: true obligatory coreference. Attempts at attributing distinct arguments to the verb and the noun will lead to neither coerced nor literal, but completely unacceptable sentences (72b and 72d):

- (72) a. *O governo promete à sociedade que, de agora em diante, vai mudar de vida* (...)
(The government promises society that, from now on, it will change [its] life)
- b. **O presidente promete mudar de vida do governo*
(The president promises to change [its] life of the government)
- c. (...) *recusando uma postura radical, a esmagadora maioria (77%) dos eleitores quer o PT participando do Governo* (...)
(Refusing a radical attitude, the overwhelming majority (77%) of voters wants the Worker’s Party to participate in the government)
- d. ??*Os eleitores recusam a postura do governo*
(The voters refuse the government’s attitude)

¹⁰ This VP is homonymous with a different sense of *confirmar* having a reportative reading:

(i) Quando perguntado, João confirmou a vantagem
(When asked, João confirmed the advantage)

This is simply a true non-obligatory coreference VP, as reporting someone else’s advantage is perfectly acceptable and does not imply any of the coercive readings proposed in this paper:

(ii) Quando perguntado, João confirmou a vantagem do Pedro
(When asked, João confirmed Pedro’s advantage)

This homonymous VP should not be confused with the one analyzed here, in which *confirmar* (confirm) means roughly “to do something in order to maintain oneself in a certain status”.

However, the former accepts no subordinate clause paraphrases and no tense mismatch but the latter does, suggesting they are not part of the same class.

Recusar a autorização (“deny authorization”) works similarly, except that the polysemy of *recusar* (“deny”/“decline”) may lead to a different interpretation of this VP if we try to impose a non-coreferential reading. In this reading, *autorização* (“authorization”) can be understood as a document that was not accepted for a certain purpose (as in 73b). However, it cannot be understood as the act of authorizing *per se*, which is why this VP, as it appears in the original corpus sentence (14b, repeated here as 73a), can be considered a case of true obligatory coreference.

- (73) a. *A ONU admitiu que os sérvios não haviam cumprido as exigências, mas recusou autorização para bombardeio.*

(The UN admitted that Serbs had not complied with the requirements, but denied authorization for bombing)

- b. *O governo recusou a autorização da ONU para bombardeio*

(The government declined the UN’s authorization for bombing)

4 Final Remarks

This pilot study departed from a description of predicative nouns and their arguments in the Brazilian Portuguese part of the Bosque Corpus, zeroing in on 75 sentences which were observed to manifest coreference between content verbs and predicative nouns in object position. These verb phrases were classified based on four properties which literature and data observation have shown to be critical in the classification of coreferential content verb phrases: types of coercion upon imposition of a non-coreferential reading, paraphrasability with subordinate light verb constructions, tense mismatch between verb and noun, and determiners.

The results show that four broader semantic classes emerge from the distribution of these properties: intention and fulfillment verbs plus action nouns, credit verbs plus dysphoric action nouns, demand and response verbs plus nouns, manifestation verbs plus psychological nouns, some of them with further subdivisions. Together, these classes cover 50 verb phrases, which seem to have the coreferential reading as the preferred one due to their lexical semantics, amounting to circa two thirds of the observations. Other 17 were considered to be verb phrases in which coreference was a possible but non-preferred reading, and 8 were considered to be phrases with preferentially coreferential readings which did not seem to fit any class, including a rare subset of true obligatory coreference.

These results point out some additional directions in relation to previous work, suggesting different syntax-semantics relations to coreference in content verb plus noun constructions. They show that both verb and noun play crucial roles in establishing whether coreference is the usual reading of a phrase, while observing semantic coercion in these sen-

tences and showing its potential as a classificatory feature. The distributional features also proved to be distinctive of classes.

Also, this seems to be the first proposal of a study on coreference between arguments of content verbs and predicative nouns in Brazilian Portuguese and the first investigation on this issue to draw its data primarily from a corpus rather than intuition alone.

This pilot study, whose limitations regarding corpus representativity and size were set out from the beginning, did not intend to provide a complete description of its object, let alone to build a theoretical model of its syntax-semantics. Given the small number of examples, around 10.7% of the data could not be neatly organized into classes, and many *hapax legomena* remain unexplored. While a class of outliers is not necessarily a flaw of this study, since irregularity is as much a feature of language as regularity, some of these eight phrases show enough regularity to be assigned a classification in a larger corpus study.

Still, this study fulfills its role of presenting a small-scale empirical investigation of this understudied phenomenon, presenting a new perspective to inform both theoretical models and further, more extensive corpora investigations into this topic while adding insights to bridge the gap between lexical and phrase semantics.

References

- AFONSO, S.; BICK, E.; HABER, R.; SANTOS, D. Floresta sintá(c)tica: a treebank for Portuguese.
- RODRIGUES, M.; ARAUJO, C. (Org.) *Proceedings of the Third International Conference on Language Resources and Evaluation* (Las Palmas de Gran Canaria Espanha, 29-31 de Maio de 2002), Paris: ELRA, 2002.
- ALBA-SALAS, J. Subject control into nominals in Romance. *SKY journal of linguistics*, v. 19, p. 9-62, 2006.
- BARROS, C. *Descrição e classificação de predicados nominais com o verbo-suporte fazer no Português do Brasil*. Doctoral thesis. São Carlos: Universidade Federal de São Carlos, Programa de Pós-graduação em linguística, 2014.
- BRUENING, B. Light verbs are just regular verbs. *University of Pennsylvania Working Papers in Linguistics*, v. 22, n. 1, p. 7, 2016.
- BUTT, M. The light verb jungle: Still hacking away. AMBERBER, M.; BAKER, B (Ed.), *Complex predicates in cross-linguistic perspective*, Cambridge: Cambridge University Press, p. 48-78, 2010.
- CALCIA, N. “Este foi outro aspecto que sofreu uma avaliação positiva”: as construções conversas fazer-sofrer, *Estudos Linguísticos*, v. 52, n. 2, p. 579-594, 2022.
- FOTOPOULOU, A.; LAPORTE, E.; NAKAMURA, T. Where Do Aspectual Variants of Light Verb Constructions Belong? *Proceedings of the 17th Workshop on Multiword Expressions*, p. 2-12, 2021.
- GROSS, G. Les constructions converses du français. *Langue et cultures*, 22. Travaux du Laboratoire de Linguistique Informatique. Librairie Droz: Genève Paris, 1989.
- GROSS, M. *Méthodes en syntaxe: régime des constructions complétives*. Paris: Hermann, 1975.
- GROSS, M. Les bases empiriques de la notion de prédicat sémantique. *Langages*, n. 63, p. 7-52, 1981.
- GROSS, M. La fonction sémantique des verbes supports. *Travaux de Linguistique*, v. 37, n. 1, p. 25-46, 1998.

- LAPORTE, E. Exemplos atestados e exemplos construídos na prática do léxico-gramática. *Revista (Con)textos Lingüísticos*, v. 2, p. 26-51, 2008.
- MEL'ČUK, I. Collocations and lexical functions. COWIE, A. (Ed.), *Phraseology: Theory, analysis, and applications*, Oxford: Oxford University Press, p. 23-53, 1998.
- MEYERS, A; REEVES, R; MACLEOD, C. NP-external arguments: A study of argument sharing in English. *Proceedings of the Workshop on Multiword Expressions: Integrating Processing*, p. 96-103, 2004.
- MOUROUNAS, M. *Nominalisation and inherent control*. Doctoral thesis, University College London, 2019.
- PETRUCK, M.; ELLSWORTH, M. Representing support verbs in framenet. *Proceedings of the 12th Workshop on Multiword Expressions*, p. 72-77, 2016.
- RASSI, A. *O verbo dar em português brasileiro: descrição, classificação e processamento automático*. Araraquara, SP: Letraria, 2023.
- RODRIGUES, R.; PICOLI, L. O modelo do Léxico-Gramática no Brasil. *Revista de Estudos da Linguagem*, v. 27, n. 2, p. 885-910, 2019.
- SAG, I.; BALDWIN, T.; BOND, F.; COPESTAKE, A.; FLICKINGER, D. Multiword expressions: A pain in the neck for NLP. *International Conference on Intelligent Text Processing and Computational Linguistics*, Berlin: Springer, p. 1-15, 2002.
- SANTOS, M. *Descrição dos predicados nominais com o verbo-suporte 'ter'*. Doctoral thesis, São Carlos: Universidade Federal de São Carlos, Programa de Pós-graduação em linguística, 2015.
- SARDINHA, T. The Brazilian Corpus. *AACL 2009 Abstract Book*, Alberta, p. 30, 2009.
- SAVARY, A. et al. The PARSEME shared task on automatic identification of verbal multiword expressions. *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)*, Valência, p. 31-47, 2017.