A STUDY OF PREPOSITIONS IN THE CONTEXT OF AUXILIARY VERB CONSTRUCTIONS

Sueli Maria COELHO*

▪ RESUMO: Neste artigo, descrevemos e analisamos construções da língua portuguesa formadas por [verbo auxiliar + preposição + infinitivo], que são aqui abordadas num viés pancrônico, pautado numa concepção cognitivista e multissistêmica de língua(gem). Nosso objetivo maior consiste em explicar por que, em algumas dessas construções, a preposição tende a ser apagada, sobretudo em contextos de oralidade, sem que tal apagamento resulte em prejuízo para a boa formação da sentença, enquanto em outras o liame preposicional deve obrigatoriamente se interpor entre o verbo auxiliar e a forma nominal de infinitivo, sob pena de se comprometer a boa formação do composto. No entorno dessa questão maior, buscamos ainda descrever que tipo de preposição ocorre nessas construções, a correlação entre o tipo de preposição e a função gramatical (tempo, aspecto, modalidade) codificada pela construção, além de identificar fatores condicionantes da presença da preposição nesses contextos. Os resultados obtidos acusam que apenas as preposições A, DE, PARA e POR podem ocorrer nesses contextos e que elas são herdadas do contexto de reanálise da construção, majoritariamente ligado a um ambiente sintático de finalidade. A opcionalidade da preposição no contexto é restrita à preposição A e condicionada por uma imbricação de fatores de natureza fonológica, prosódica e semântica.

▪ PALAVRAS-CHAVE: Preposição. Construção de verbo auxiliar de incidência indireta. Aspecto verbal. Tempo. Modalidade.

Description of the object of the study, inquiry, and methodology

This article is dedicated to the study of auxiliary verb constructions of indirect incidence, meaning those whose structure presents a prepositional link between the morphemic verb and the nominal infinitive form: [V1 + preposition + V2$_{infinitive}$]. According to Joao de Almeida (1980, p.28, our translation), "[…] the phrase 'auxiliary plus infinitive' dominates the panorama of verbal periphrasis in the Portuguese language, and is, therefore, where the auxiliary process becomes more constant and more varied." However, despite the phenomenon's productivity and the immense wealth of studies

* Universidade Federal de Minas Gerais (UFMG). Belo Horizonte – MG - Brazil. sucoelho@ufmg.br. ORCID: 0000-0003-4021-0339

on the auxiliary verb in Romance languages, there are practically no broader works devoted to investigating it under the terms of our proposal,[1] which, therefore, justifies our investment in this area. Our portion in the scope of Portuguese language auxiliarity is thus restricted to the constructions [V1 + preposition + V2$_{infinitive}$] and seeks to answer four key questions to be discussed in separate sections throughout this text: (i) Which prepositions may occur in this context?; (ii) Is there specificity of function related to the type of prepositional link or do the same prepositions participate in the coding of different functionalities?; (iii) Why can some prepositions be removed from the construction without prejudice to their good formation, while others cannot?; (iv) What factor(s) determine(s) the selection of the preposition in this context?

To try to answer these questions, we quantitatively and qualitatively analyzed a set of 802 (eight hundred and two) constructions gathered from the Tycho Brahe historical *corpus* (http://www.tycho.iel.unicamp.br/corpus/). The collection covered the period between the fifteenth and twentieth centuries and was conducted in accordance with the labeling system of the *corpus* itself: we opted for the type of search that would allow us to graphically construct the query and, aiming to obtain the most comprehensive sample of the language, we determined that the search would be processed in all labeled texts, thus covering a greater diversity of textual genres. The label constructed for the query was: [verb >> preposition >> infinitive verb].

Once in possession of the search results, we selected the first 150 (one hundred and fifty) occurrences for each century to compose our *corpus*. However, we found that not all of them were auxiliary verb constructions, leading us to eliminate 98 (ninety-eight) occurrences that were only juxtaposed forms, such as, *tem para agradecer (one has to be thankful)* and *é de cortar (it is to cut)*. Such occurrences were eliminated from the quantification but were considered in the qualitative analysis to help us identify possible reanalysis contexts of the auxiliary verb, for which we also relied on data collected in the *Portuguese Corpus* (https://www.corpusdoportugues.org/).

Aware of the methodological limitation that the work with a *corpus* imposes on the researcher, especially with regard to the possibility of not allowing him/her to identify constructions that his/her speaker-listener intuition believes to exist, but which are not documented, we endorse Santana's understanding (2010, p.132, our translation) that "[…] the most effective methodological procedure for overcoming this type of limitation of data emanating from corpora is to mix data collection with the construction of examples based on intuition." In addition, given that the synchronic outline established for this study did not allow us to include orality data at the risk of

---

[1] The only work of its kind that we could identify was that of Sousa (2011), who, guided by Prof. Dr. Heloísa Maria Moreira Lima de Almeida Salles (UnB), presents a propaedeutic study of the topic in which we are engaged in her final work on the subject Course Design. Supported by the generative theoretical framework, she explored the hypothesis that prepositions that can be removed in the context of verbal periphrases are not predicators, but rather markers of an abstract case. However, given the preliminary nature of the work, the author recognizes that "there is much to go into more deeply" and reaffirms the "[…] need for a study dedicated to verbal periphrases, especially those of the infinitive, which would cover all contexts of the use of prepositions, not only the case of the aspect." (SOUZA, 2011, p.32, our translation).

sacrificing the *corpus* balance, our methodological option was to fill this gap with the construction of illustrative examples of real situations of oral interaction. Thus, although our quantitative analysis is anchored in data collected in *corpora*, our qualitative analysis does not shy away from considering intuition data properly licensed by Portuguese grammar. Therefore, in this text, the data extracted from *corpora* are properly referenced and those that need not be referenced should be understood as intuition data, often preferred here because they allow us to illustrate our argument in a more didactic way. That said, we shall now begin the task of analyzing our first question.

Prepositions in the syntactic environment of auxiliary verb construction

Given that auxiliarity results from a grammaticalization process[2] in such a way that it occupies the V1 position, and the prepositions are already in their grammatical forms, our expectation was that not all types of prepositional links could occur in context, a hypothesis supported by the restrictions arising from the grammaticalization stage of the construction and then confirmed when analyzing the data collected in the historical *corpus*. With the intention of identifying which prepositions can occur in the auxiliary constructions with the infinitive, we analyzed the following statements:

(1) Maria *começou **a** fazer* dieta e, desde então, não sai mais para jantar com os amigos.
(Maria started on a diet and, since then, she has not gone out to dinner with friends.)

(2) O menino tem tanto medo de injeção, que já *dispara **a** gritar* só de ver alguém vestido de branco.
(The boy is so afraid of a shot that he starts to scream just seeing someone dressed in white.)

(3) Alguém *chegou **a** falar* pessoalmente com o homenageado, para confirmar sua presença no evento?
(Did anyone speak personally with the honoree to confirm their presence at the event?)

(4) Com a mudança de estação, a criançada *deu **para** adoecer* e quase todos os dias é um que perde aula.
(With the change of season, the children began to get sick and almost every day one of them misses classes.)

(5) O melhor é se vacinar e se prevenir; afinal, não *dá **para** adoecer* todo dia, não é mesmo?
(It's best to get vaccinated as a precaution; after all, you can't get sick every day, can you?)

---

2   For the purposes of this study, the grammaticalization process that involves the formation of auxiliary verbs is conceived in the words of Kurylowicz (1965, p.52): "Grammaticalization consists of the increase in the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status."

(6) Os convidados *estão **para** chegar* a qualquer momento.
(The guests are due to arrive at any time.)

(7) *Acabei **de** ler* o livro que o professor recomendou.
(I've just read the book that the teacher had recommended.)

(8) a. Ninguém *tem **de** assumir* a culpa só para poupá-lo da punição.
(No one has to take the blame just to save him from being punished.)
b. Ninguém *tem **que** assumir* a culpa só para poupá-lo da punição.
(No one has to take the blame just to save him from being punished.)

(9) Seu quarto *ficou **por** arrumar* a semana toda; hoje você precisa dar um jeito nele.
(Your room has been a mess all week; today you need to clean it.)

An analysis of the constructions highlighted in the contexts of statements (1) to (9) reveals that the list of prepositions that occur in the contemporary auxiliary verb constructions is restricted to four forms, as we could see from the analysis of the 802 data items collected diachronically: A, DE, PARA, and POR[3]. Of this quartet, the most productive in the *corpus* is the preposition DE (93.7%), followed by A (3.8%), PARA (2.2%), and POR (0.3%). This set consists of the most emblematic examples of the category, except for the preposition EM[4], recorded by Travaglia (1985, p.217) as a linguistic expedient adopted by the speaker to "indicate something not yet accomplished", but which expresses "willingness or intention":

(10) "Meu irmão *esteve **em** pedir*-lhe o terno emprestado, mas ficou com vergonha de fazê-lo". ("My brother was about ask if he could borrow your suit, but he was ashamed to do it.") (TRAVAGLIA, 1985, p.218)

Despite the author's claim that the construction [ESTAR + EM + INFINITIVE] is not a variant form of the construction [ESTAR + POR + INFINITIVE[5]], as Dias (1970) proposes, we understand these constructions are, indeed, variant forms and we argue that the construction [ESTAR + PARA + INFINITIVE] also integrates, in the context presented by the author, the set of linguistic variants that are suitable to mark the imminence of the action expressed by the nominal form of the infinitive; therefore, a temporal notion, since, unlike aspect, it has a deictic anchor. In the wake of the indirect incidence constructions that have a relational auxiliary verb (or link, in traditional terms), we identified those formed by [FICAR + POR + INFINITIVE], which mark durative aspect, as exemplified in (9). In the view of Travaglia (1985, p.242, our translation),

---

[3] These Portuguese prepositions correspond to the following English prepositions: TO (A/PARA), OF (DE), and BY (POR). In the English language, however, they do not occur in the same syntactic context of auxiliary verbs relative to the phenomenon we are studying, as they do in Portuguese.

[4] This Portuguese preposition corresponds to the English preposition IN.

[5] In the words of Travaglia (1985, p.302, our translation), "[…] the preposition **por** placed before the infinitive verb is the main, if not the only, one responsible for the expression of the unstarted aspect in Portuguese: (1370) Ainda há vários espécimes **por cataloger.**" (There are still several specimens to be catalogued).

[…] when the period of permanence in the situation is not in any way specified in the sentence or in the context, we have the logical implication that what "was or will remain to be done" "was or is to be done". This implication leads to interpret the verb "ficar" (stay) as equivalent to "continuar" (continue), which can lead us to an aspectual analysis similar to the one we have for "continuar + por + fazer"[6]

Having identified the four prepositions likely to occur in the constructions with which we are concerned, we proceed to our second question, whose answer comprises the coding of three functional categories – time, aspect, and modality – as more fully described in the next section.

Correlation between prepositional link and functional construction coding

Statements (1) to (9) indicate that the distinct functionality of auxiliary verb constructions is not exclusively determined by the type of prepositional link that appears in the compound, given that the same preposition can integrate constructions with different functionalities. However, it is important to note the fact that the functionality of the construction emerges from the compound does not imply there is no specialization of functions among the four prepositional connectors that may occur in this context. For example, the preposition A, albeit productive in this syntactic environment, does not appear in auxiliary verb constructions that encode temporality or modality. Nevertheless, it has proved to be the most productive preposition in aspectual constructions, especially inceptive (cf. 1) and durative (cf. 2) constructions. As Travaglia (1985, p.301, our translation) understands it, "[…] the preposition A, when used with the infinitive, seems to be responsible for imperfective, cursive, unfinished, and durative aspects." Furthermore, according to the author, the removal of the preposition in a construction, or even its replacement by another, interferes significantly in the expression of aspect, which seems to foreshadow the predominant role of this class in coding the aspectual category,[7] as illustrated below, using some of the data extracted from the distinguished linguist's work (TRAVAGLIA, 1985, p.301-302):

---

[6] Original: "*quando não se especifica de alguma forma, na frase ou no contexto, o período de permanência na situação, temos a implicação lógica de que o que 'ficou ou ficará por fazer' 'estava ou está por fazer'. Esta implicação leva a interpretar o verbo 'ficar' como equivalente a 'continuar', o que pode nos levar a uma análise aspectual semelhante à que temos para 'continuar + por + fazer'*".

[7] The preponderant role of the preposition in the coding of grammatical functions can be further attested by constructions with the auxiliary verb *deixar*, for example, in which, when added to the features [± subject correspondence], the presence of the preposition encodes a semantic notion of cessation (*Deixei **de** comer* o doce), while its absence results in a causative construction (*Deixei-o comer* o doce).

(11) a. "Vou **a ler** o livro pelo caminho". (imperfectivo, cursivo, não-acabado, durativo) ("I'm reading the book along the way." [imperfective, cursive, unfinished, durative])

b. "**Vou ler** o livro pelo caminho". (futuro) ("I will read the read along the way. [future]).

(12) a. "Vou **a perguntar** por meu filho". (imperfectivo, não-acabado, iterativo) ("I'm repeatedly asking about my son". [imperfective, unfinished, iterative])

b. "**Vou perguntar** por meu filho". (futuro) ("I will ask about my son". [future])

(13) a. **Acabou de comprar** um vestido simples. (terminativo) ("She's just bought a plain dress". [terminative])

b. **Acabou por comprar** um vestido simples. (resultado final ou consequência; a noção aspectual perfectiva é denotada pelo pretérito perfeito) ("She ended up buying a plain dress". [final result or consequence; the perfective aspectual notion is denoted by the past perfect])

A comparative analysis of the contexts of (13), together with those of (11) and (12), implies that some prepositions have the property of coding aspects, while others do not. That being so, if the prepositions A and DE are determinants for aspectual marking, given that their suppression in the context of the construction nullifies the expression of this semantic category, as illustrated in (11) and (12), the same could not be said of the preposition POR, since, in Travaglia's (1985) understanding, the aspectual notion of the construction in which it participates stems from the auxiliary tense and mood marking. Almeida (1980, p.28, our translation) had a similar view, for whom "[…] the periphrasis [*acabam por fugir (*ends up escaping)] does not exactly translate into duration, but rather as a result, a consequence of what must have been considered before." The assumption that not every preposition that occurs in the context of support has the potential to encode aspect is evidently confirmed when it comes to the connotation of phase aspect, but not of other aspectual notions, as we will now examine. Barroso (1994), who exhaustively described the aspectual category in contemporary Portuguese, separating it into seven subcategories, considered that (13.b.) encodes *placement aspect*. In the understanding of this scholar of European Portuguese:

> this aspectual category [...] marks the relationship of one action with another (or other) action (s) in the context.
> [...]
> The placement comprises three subcategories and all are represented peripherally in the contemporary Portuguese linguistic norm, namely: l. alignment (or order), 2. resulting arrangement, and 3. demarcation. (BARROSO, 1994, p.137).[8]

The construction highlighted in (13.b.) would correspond to the aspectual subcategory of alignment (or order) in Barroso's categorization (1994), according to whose principle

> In Portuguese, however, only the 'alignment' of verbal action at its beginning and [...] at its end is performed [...] peripherally. The signifiers (= periphrases) that express this aspect in the Portuguese linguistic norm are **começar + por + infinitive** (or **começar + gerund**), at the beginning; **acabar + por + infinitive** (or **acabar + gerund**) and **terminar + por + infinitive** (or **terminar + gerund**), at the end (BARROSO, 1994, p.137-138, our translation).[9]

Thus, it can be seen that the behavior of the preposition POR in (13) is very similar to that of the prepositions A and DE in contexts (11) and (12), since they can also alternate with the nominal form of the gerund, as described by Barroso (1994) and recorded by Almeida (1980), which leads us to assume that the three prepositions can contribute to the coding of the aspect. Considering, however, that grammaticalization processes necessarily involve specialization of functions (Cf. HOPPER, 1971), it is possible to understand why some forms are more salient than others in certain uses, as seems to be the case of A and DE in relation to POR in the contexts described above. Moreover, even in relation to forms that denote the same phase, there are distinct semantic nuances arising from the degree of cohesion of the forms, as well as from the original sense of the auxiliary, or even from the semantics of the construction's preposition, as observed in the statements below, extracted from Almeida (1980, p.44):

(14) "[...] e como se tivesse explicações a dar-me *começou por dizer* que se chamava Kraus [...]". (and as if he had explanations to give me, he started by saying that his name was Kraus.)

(15) "E os filhos já *começaram a reproduzir* o gesto hereditário". ("And the children have already started to reproduce the hereditary gesture".)

(16) "Agora sim, *começará* o nosso telégrafo *a trabalhar*". ("Now, our telegraph will begin to work".)

In view of these examples, Almeida (1980, p.43, our translation) considered that:

---

linguística portuguesa contemporânea, a saber: l. alinhamento (ou ordem), 2. disposição resultante e 3. demarcação." (BARROSO, 1994, p.137)

[9] Original: "[...] *uma acção pode alinhar-se no seu começo, no meio, ou no seu termo. Em português, porém, só o 'alinhamento' da acção verbal no seu começo [...] e no seu termo [...] se encontra realizado perifrasticamente. Os significantes (= perífrases) que expressam este(s) valor(es) aspectual(ais) na norma linguística portuguesa são começar + por + infinitivo (ou começar + gerúndio), para o começo; acabar + por + infinitivo (ou acabar + gerúndio) e terminar + por + infinitivo (ou terminar + gerúndio), para o fim*." (BARROSO, 1994, p.137-138).

[…] if inception is common to the two constructions, it is expressed in each one for different purposes. [...] It is necessary, specifically, to observe the greater integration between the constituent elements of the periphrasis with the preposition *a*, while the preposition is configured as an obstacle between the auxiliary and the infinitive dynamics, highlighting the notion of difficulty.[10]

Almeida's interpretation (1980) finds its echo in the categorization proposed by Barroso (1994), who conceives the construction [COMEÇAR + POR + INFINITIVE] as a mechanism that allows the speaker to establish an alignment or ordering of actions in relation to its chronology. In the case of statement (14), although he owed other explanations to the speaker, the character started by telling him his name, perhaps because he considered this action to be the simplest and most understandable of all that he was responsible for explaining.

In the constructions of (15) and (16), the speaker does not notice any intention of aligning actions, but only a fundamental reference to the initial phase of the verbal process expressed by the nominal infinitive form, with no concern for its later stages. There are cases, however, in which the preposition A also includes constructions that connote more than one aspectual notion, one of which is linked to inception and the other to the gradual intensity of the event, which culminates in its iteration:

(17) Desde que *pegou a andar* em más companhias, o adolescente não se relaciona mais tão bem com os pais. (Ever since he started keeping bad company, the teenager does not get along very well with his parents.)

(18) Maria ficou bêbada e *destampou a falar* palavrões, constrangendo a todos. (Maria got drunk and let loose with a string of profanity, embarrassing everyone.)

(19) O homem ficou nervoso e *garrou a xingar* todos que se aproximavam dele. (The man became angry and started to curse everyone who came near him.)

(20) De repente, o homem *danou a jogar pedras* na multidão. (Suddenly, the man began to throw rocks at the crowd.)

All the constructions highlighted above encode more than one aspectual notion, namely, the initial phase of the event expressed by the nominal infinitive form – which was triggered by a change of state (inchoation) – its duration, and consequent reiteration over a period of time. The difference in these constructions in relation to those illustrated in (15) and (16), which mark only inception, resides primarily in the lexical nature of the auxiliary, more semantically emptied in one group of constructions than in the

---

[10] Original: "[…] *se a inceptividade é comum às duas construções, ela se expressa em cada uma com propósitos diferentes. [...] Há que, especificamente, observar a maior integração entre os elementos constitutivos da perífrase com a preposição a, enquanto que a preposição por se configura como um obstáculo entre a dinâmica do auxiliar e a do infinitivo, pondo em relevo a noção de dificuldade.*"(ALMEIDA, 1980, p.43).

other. According to Almeida (1980, p.42), it is necessary to distinguish "two groups of peripheral expressions", which are inceptive in Portuguese, namely:

> 1st) that of periphrases in which the inceptive idea comes from the auxiliary itself;
>
> 2nd) that of periphrases in which, having reduced the auxiliary's lexical meaning in favor of the grammatical meaning, the inceptive idea is part of the "auxiliary, preposition, and infinitive" set (ALMEIDA, 1980, p.42, our translation).[11]

Based on Almeida's generalization (1980) concerning the difference in the semantic load of the auxiliary – which, undoubtedly, indicates different degrees of grammaticalization between the two groups of constructions – and the specificities identified among them in the coding of the aspectual function, it is possible to infer that the periphrases of the first group are less grammaticalized than those of the second, which can be attested by the possibility of inserting the SN [*nosso telégrafo* (our telegraph)] between the auxiliary and the infinitive (16) and by the overlapping of grammatical functions related to the coding of more than one aspectual notion in constructions (17) to (20), which translate both stage aspect and extension aspect.[12]

Similar functional behavior could be observed in the construction [DAR + PARA + INFINITIVE], which, in addition to the semantic emptying of the auxiliary, shares with the constructions highlighted in statements (17) to (20) the property of condensing more than one aspectual connotation, as illustrated in (04), repeated here under new numbering for the convenience of the reader, and in the following statements (22) to (23):

(21) Com a mudança de estação, a criançada *deu **para** adoecer* e quase todos os dias é um que perde aula. (With the change of season, the children began to get sick and almost every day one of them misses class;)

(22) Joaquim *deu **para** beber* depois que perdeu o emprego. (Joaquim started drinking after he lost his job.)

(23) Alguns funcionários *deram **para** reclamar* da obrigatoriedade do ponto eletrônico. (Some employees started complained about the mandatory electronic timecard.)

---

[11] Original: "*1º) o das perífrases em que a idéia inceptiva parte do próprio auxiliar; 2º) o de perífrases em que, diminuído o auxiliar de significação léxica, em favor da significação gramatical, a idéia inceptiva parte do conjunto "auxiliar, preposição e infinitive*." (ALMEIDA, 1980, p.42).

[12] Almeida (1980, p.39-40, our translation) proposed breaking down the aspect category into two basic types: (i) the *lato sensu* aspect, which brings together the perfective/imperfective duality, and (ii) the *stricto sensu* aspect, "[…] which is fundamentally characterized by the notions of inception, cursivity, termination, punctuality, duration, iteration, and globality. And these notions can be opposed in two distinct groups, which, for the sake of convenience, we will call 'phase aspects' and 'extension aspects'. In the first group, we include inception, cursivity, and termination; in the second, duration, punctuality, and iteration, with globality representing the neutralization of the 'durative-punctual' opposition'."

Considering that, since Latin, the prepositions A and PARA are competing ways to translate direction or movement, it is predictable that both words present some functional similarity as well as some specialization when integrating auxiliary verb constructions, which is confirmed empirically, as we have shown. Our interpretation of such an occurrence shows that there are preposition specializations for certain functions, which probably arise from the semantic-cognitive specificities of the forms that are or are not preserved in the course of their grammaticalization process in the language and from their diachronic history, whose features determine the competition between some forms in certain contexts of use.[13] Accordingly, our analysis indicated that, while the preposition A can be combined with auxiliary verbs that still preserve their original semantic value (Cf. (15) and (16)) and with those that are emptied of it (Cf. (17) to (20)), resulting in different functionalities relative to the strict coding of stage aspect or the conjunction of stage aspect with extension aspect, the preposition PARA only combines with auxiliary verbs that are emptied of their original semantic value in the course of the grammaticalization process (Cf. (21) to (23)), thus connoting more than one aspectual notion.[14] This restriction suggests that the preposition PARA carries more semantic features than the preposition A, a hypothesis based on the thesis of phonetic erosion, in accordance with Heine, Claudi and Hünnemeyer (1991), and on the parameter of paradigmatic integrity proposed by Lehmann (1995).

Another preposition identified in the syntactic context of constructions that have the verb DAR as an auxiliary was DE:

(24) Desde que descobriu a verdade sobre sua origem, a jovem *deu **de** ficar* triste pelos cantos. (The young woman has been moping around ever since she found out the truth about her origin).

(25) De uns tempos para cá, minha mãe *deu **de** falar* sozinha. (My mother has been talking to herself for lately.)

(26) Ele agora *deu **de** ficar enviando* flores para a namorada; deve estar apaixonado. (He suddenly started sending flowers to his girlfriend; he must be in love.)

According to our analysis, the construction [DAR + DE + INFINITIVE] is a variant of the construction [DAR + PARA + INFINITIVE], since the exchange of one form by the other does not change the true value of the statements in these contexts. Furthermore, these constructions also seem to be in the process of linguistic variation with the one highlighted in (26), in which the nominal form of the infinitive is replaced by the auxiliary *ficar + gerund*, a device that seems to consist of an attempt to reinforce the coding of extension aspect, as these data illustrate:

---

[13] As Said Ali reported (1966 [1921], p.211, our translation), "[…] the [preposition *a*] in Latin is used to enunciate the concept of direction or movement to some point, of approach and final junction of one thing to another. This same meaning still lives in our preposition *a*, despite the competition with *para*, which sometimes restricts its function."

[14] According to Almeida (1980, p.68, our translation), this "[…] verbal periphrasis offers an interpretation in favor of inception and iteration, naturally with the nuance of whimsy or quirkiness or a strange or unusual attitude."

(25)   a. De uns tempos para cá, minha mãe *deu **de** falar* sozinha.
       b. De uns tempos para cá, minha mãe *deu **para** falar* sozinha.
       c. De uns tempos para cá, minha mãe *deu **de** ficar falando* sozinha.
       d. De uns tempos para cá, minha mãe *deu **para** ficar falando* sozinha.
          (My mother has been talking to herself lately.)

An idiosyncrasy was also identified when it came to the construction [DAR + PARA + INFINITIVE]. In this specific construction, multifunctionality is not only of the preposition, but of the entire set, which, in addition to coding aspect, may also denote modality, as illustrated in (04) and in the following contexts:

(27)   Não *dá* mais ***para** acreditar* em mudanças. (You can't believe in changes anymore.)
(28)   Hoje não *deu **para** ir* ao banco; farei isso amanhã. (I couldn't go to the bank today; I'll do it tomorrow.)
(29)   Será que *dá **para** fingir* que não aconteceu nada? (Is it possible to pretend that nothing has happened?)

The constructions highlighted in these contexts denote epistemic modality, as they express the attitude of the speaker with regard to the value of (im)possibility that is imputed upon the propositional content of the statement.[15] However, there is a syntactic difference regarding the selection of the subject between these two apparently identical constructions from the point of view of form: while the aspectual construction (Cf. (30)) requires a determinate subject, the modal construction (Cf. (31)) demands an indeterminate subject:

(30)   Aquele menino *deu **p(a)ra** comer* doce no café da manhã. (That boy started eating candy for breakfast.)
(31)   Não *dá **p(a)ra** comer* doce no café da manhã! (One can't eat candy for breakfast!)

As we have shown, the preposition PARA only participates in the coding of modality in constructions with the auxiliary DAR and, even so, when the subject is indeterminate.[16] It is, therefore, a very specific contextual constraint, which is replicated with the auxiliary ESTAR, in example (06) – repeated here for the convenience of the

---

[15]   It can also be avowed that such constructions encode an agent-oriented modality, in the sense that they evoke the possibility that conditions internal or external to the agent are acting on the action expressed by the main predicate. However, such ambiguity is not a problem, since, according to Bybee, Perkins, and Pagliuca (1994, p.195), "[…] the fact that some of the English modal auxiliary verbs have both agent-oriented or root meanings and epistemic ones is well known [...] It is clear that the epistemic senses develop later than, and out of, the agent-oriented senses."

[16]   The possibility that this construction does not connote aspect when combined with a determined subject is restricted to cases in which it is used metaphorically and is not endowed with the semantic role of agent, denoting rather an idea of (in) sufficiency, as in *Time was not enough for us to do what we intended to,* for example.

reader – in which the construction [ESTAR + PARA + INFINITIVE] encodes a temporal notion of near future, which translates into the imminence of the action expressed by the nominal form of infinitive:

> (06) Os convidados *estão **para** chegar* a qualquer momento. (The guests are due to arrive at any time now.)

With regard to the multifunctionality and functional specialization of prepositions, the data revealed that the most productive preposition in Portuguese for the coding modality is DE, already identified in this function since the 15th century, according to the occurrences in our *corpus* extracted from the Tycho Brahe database (http://www.tycho.iel.unicamp.br/corpus/):

> (32) "E Brás *há **de** endoudecer* pois Deus não *há **de** querer* que eu nada faça de mim." ("And Brás will go crazy, because God will not want me to do nothing of myself".) (Gil Vicente, 15th century)
>
> (33) "Mas porque a mãe sabe o fim que *hão **de** dar* a esta criança, muitas vezes, quando se sente prenhe, mata-a dentro da barriga [...]" ("But because the mother knows what they will do to her child, when she finds out she is pregnant, she often provokes a miscarriage".) (Pero Magalhães de Gândavo, 16th century)
>
> (34) "De onde *havemos **de** tirar* êste dinheiro, estes navios, esta gente de mar eguerra que *havemos **de** dar* a França todos os anos?" ("From where will we take this money, these ships, these men of the sea and war that we must give to France every year"?) (Letters by Fr. Antônio Vieira, 17th century)
>
> (35) "Mas parece-me que algumas scenas *tem **de** ganhar* muito em ser abreviadas [...]" ("But it seems to me that some scenes will be improved when they are shortened".) (Letters by Almeida Garrett, 18th century)
>
> (36) "*Tenho **de** fazer* reparos urgentes no sobrado deste engenho." ("I have to make urgent repairs to the Master's house on this plantation".) (Brazilian letters, 19th/20th centuries)

Before proceeding to the analysis of the modal constructions highlighted in the above-cited contexts, it should be noted that, diachronically, such constructions are, for the most part, more productive than the aspectual ones. Additionally, the construction of modal auxiliaries of indirect incidence in the Portuguese language is primarily made by using the auxiliaries HAVER and TER,[17] since the other possibility, the construction [DAR + PARA + INFINITIVE], is much less productive, perhaps because it also connotes aspect, observing the restrictions already discussed. When treating

---

[17] In Spanish, such periphrases include the verb deber, which can appear in modality constructions both in syntactic contexts of direct incidence, i.e., followed only by the infinitive <deber + infinitive>, and in the context of indirect incidence, i.e., followed by preposition <deber de + infinitive> (Cf. BOSQUE; DEMONTE, 1999).

modal periphrases, more specifically those formed by [HAVER + PREPOSITION + INFINITIVE], Almeida (1980, p.142, our translation) stated that:

> […] of the various periphrases in the language, for the expression of obligation, this is the one that reveals the greatest number of semantic features, and, consequently, the greatest extent of use. It lends itself to a greater subjective participation in the speaking subject's obligation process, due to what is easily linked to other moral attitudes other than mandatory ones, such as volition or desire, possibility, effort, and even an order or request[18].

The range of uses of [HAVER + DE + INFINITIVE] that have discussed can be historically attested by the contexts of (32) to (34), in which its use can take on a character of imprecation or conviction (Cf. (32)), even moral obligation (Cf. (34)). The construction [TER + DE + INFINITIVE], in turn,

> […] fundamentally serves to express the need, or rather, material obligation and logical obligation, appearing less as an expression of the moral obligation. Representing an [...] external obligation, it somehow denounces a passive character and corresponds to the phrase being obliged to. As it involves the notion of the inevitable, it may also correspond to *ser destinado a.* (ALMEIDA, 1980, p.148, our translation).[19]

If the expression of need is absolutely made in the context of (36), the same cannot be said of the construction highlighted in (35), whose meaning seems to denote a possibility of gain in the face of the abbreviation of some scenes rather than a logical obligation itself. This range of modal values is most likely due to the grammatical stage of the forms. Our analysis of the diachronic data collected from the Tycho Brahe Corpus suggests that the grammaticalization process of the verb TER in the modal auxiliary is correlated with the grammaticalization process of the verb HAVER as a marker of future. It would, therefore, be an embedded change through which each auxiliary would have specialized in the marking of a specific type of modality in the course of the history of the language, although diachronically, these functions were not always so clearly demarcated. An analysis of the data above shows that (33) and (34) are ambiguous

---

[18] Original: "*das diversas perífrases existentes na língua, para a expressão da obrigatoriedade, esta é a que revela maior número de traços sêmicos, e, consequentemente, maior extensão de uso. Presta-se a uma maior participação subjetiva no processo de obrigação do sujeito falante, em razão do que facilmente se liga a outras atitudes morais que não a de obrigatoriedade, tais como a de volição ou desejo, a de possibilidade, a de esforço e mesmo a de ordem ou pedido.*" (ALMEIDA, 1980, p.142)

[19] Original: "[…] *serve fundamentalmente à expressão da necessidade, ou melhor, da obrigação material e da obrigação lógica, aparecendo menos como expressão da obrigação moral. Representando [...] uma obrigação externa, denuncia de certa forma um caráter passivo e corresponde à locução ser obrigado a. Pode ainda, tendo o envolvimento da noção do inevitável, corresponder a ser destinado a.*" (ALMEIDA, 1980, p.148).

contexts in which both a temporal reading of the future and an epistemic (Cf. (33)) or deontic (Cf. (34)) reading are also legitimate.[20] Furthermore, the diachronic analysis undertaken implies that the context of reanalysis of the construction whose auxiliary is the verb TER is that in which its verbal complement was a transitive noun that, in front of the verb in a relative sentence, was selected as a complement in the phrase [DE + INFINITIVE], as illustrated by these data:

(37) "Paréçe que menos autoridádes bastávam pera os hómens sentirem quanta **obrigaçám** *tem de ensinár* a doutrina de Cristo" ("It seems that less authority was enough for the men to understand their obligation to teach the Christian doctrines.") (15th century)

(38) "[...] inda que do **desejo** que *tenho de servir* aVossa Mercê" ("[...] even if I wish to serve you, sir") (17th century)

(39) "[...] conheço a **obrigação** que todos *temos de empenhar* o sangue nesta ocasião." ("[...] I understand the obligation that all of us have to give your heart and soul on this occasion".) (17th century)

Drawing on the less-than-secure resource of interpreting the present in light of the past, it seems possible to infer that, as the illocutionary content of the noun was incorporated into the construction, the complement of the verb TER was eliminated as redundant and the then-lexical verb was reanalyzed as an auxiliary verb. According to what is perceived by the notional content of the nouns to which the phrase [DE + INFINITIVE] serves as a complement, at first, it could be both a deontic and an epistemic modality, as illustrated by the data from (35) and (36). With the advance of the grammaticalization process of the auxiliary verbs HAVER and TER, which have long been competing forms in the language, there was a specialization of uses: the construction [HAVER + DE + INFINITIVE] began to translate the desire for something to become effective, while the construction [TER + DE + INFINITIVE] was consolidated to mark an imposition. According to Said Ali (2008, p.142, our translation), "[…] this second form has the advantage of more precisely expressing the imperative need, the act to be performed regardless of the will. It is, however (*sic*) this application of the verb *to have* a neologism enshrined in the literary language from the eighteenth century onwards." Notwithstanding this scholar of the history of the Portuguese language, considering that the origin of the so called neologism is not clear, he may "[…] appear to have come from the habit of omitting a noun that intelligence would complete without cost, v. g. *(obrigação) tenho de ir*. A very simple

---

20 Even in Spanish there are still ambiguous contexts and, according to Bosque and Demonte (1999, p.3338, our translation), the classification of periphrases in this language cannot be done in an airtight manner. "Concerning this classification we might say, first, that a number of periphrases could appear in more than one separated, as we can see. Thus, < *ir a* + infinitive> is sometimes aspectual and other times modal or temporal, and <*haber de* + infinitive > has sometimes a future tense character and others a modal value of obligation. Second, we do not have tense periphrasis per se, even though <*haber de* + infinitive> and < *ir a* + infinitive> present, sometimes, this modal or aspectual character, what confirms that the notions of aspect, tense, and mood not always have clear boundaries."

explanation, but without facts to confirm it." (SAID ALI, 2008, p.142, our translation), we consider that the data arranged from (37) to (39) are supporting sources of the etymology we have proposed, which is also supported by Almeida (1980, p.151, our translation): "It seems appropriate to glimpse the early construction of the periphrasis there, the stage prior to the process of its grammaticalization, which naturally could only have happened after the ellipse of the noun, which left its peripheral expression on the semantic content."

Although we have not identified such competition in the diachronic data collected in the *corpus*, it is known that, in contemporary times, the preposition DE tends to compete in a very productive way with the connective QUE in the constructions of auxiliary verbs of modality, especially when the auxiliary is the verb TER, as illustrated in (8), renumbered here:

(40) a. Ninguém *tem **de** assumir* a culpa só para poupá-lo da punição.
      b. Ninguém *tem **que** assumir* a culpa só para poupá-lo da punição.
         (No one has to take the blame only to save him from punishment.)

Such competition is equally fruitful in Spanish and, as in Portuguese, it also occurs only with the auxiliary verbs *haber* and *tener*. According to Bosque and Demonte (1999, p.3338, our translation),

> […] the modal [periphrasis] only presents two cases with preposition (*<haber de* + infinitive> and <*deber de* + infinitive>). The conjunctive element *que* in <{*tener que/haber que*} + infinitive> seems to come from the relative *que* from sequences such as *Tengo cosas que hacer > Tengo que hacer (cosas), Hay cosas que hacer > Hay que hacer (cosas)* (I have things to do > I have to do (things), There are things to do > We have to do (things)[21].

This also seems to be the understanding of most of our grammarians regarding the origin of the construction [TER + QUE + INFINITIVE] in the Portuguese language, although many normativists, including Napoleão Mendes de Almeida (1969, p.226, our translation), do not believe that both forms translate the same true value, prescribing specific contexts of use to each one:

> […] when the second verb is intransitive or, even when there is no antecedent, neither expressed nor hidden, it will be better to use *de*, since the idea is always of necessity, of obligation. Let us observe Vieira's

---

[21] Original: "[…] *las [perífrasis] modales sólo presentan dos casos com preposición (<haber de + infinitive> y <deber de + infinitivo>). El nexo conjuntivo que de <{tener que/haber que} + infinitivo> parece proceder del relativo que a través de secuencias como Tengo cosas que hacer > Tengo que hacer (cosas), Hay cosas que hacer > Hay que hacer (cosas)*." (BOSQUE; DEMONTE, 1999, p.3338).

correction: "... para se conhecerem os amigos, haviam os homens *de* morrer primeiro e daí a algum tempo ressuscitar" ("to get to know their friends, the men had to die first and then, some time later, resuscitate"). *Haviam* is used for *tinham*, but, as the second verb is intransitive (*morrer*), Vieira employs, with the meticulousness of those who know the language very well, the preposition *de* instead of the pronoun *que*, to which no function would fit in the sentence.[22]

However, this is not our understanding of the linguistic fact, since, endorsing the words of João de Almeida (1980, p.152, our translation), we believe that, "[…] despite the value of such authoritative opinions, the truth is that the survey that we did [...] would not in any way allow us to deny that the expression TER QUE + INFINITIVE has the same peripheral value recognized in TER DE + INFINITIVE."[23]. Furthermore,

> […] there is little point in arguing that its more common function is that of a relative pronoun or conjunction. In the periphrasis of obligation, what remains a subordinate connective, but characteristically as a preposition, represents yet another shift in the language, of a system in motion, similarly to what has already occurred with *durante*, *mediante, mal*, etc.
>
> The opposition *ter de/ter que* offers to the speaker of the language, for the modal expression of obligation, a very expressive resource based on the diversity of the articulation point of the phoneme /d/, dental occlusive, compared to /k/, velar occlusive. (ALMEIDA, 1980, p.152, our translation).[24]

In addition to participating very productively in coding the modality in auxiliary verb constructions, as described here, the preposition DE also acts in aspectual constructions, essentially those that denote terminative aspect, along the lines of what was exemplified in (7), reproduced here under new numbering:

---

[22] Original: "[…] *quando o segundo verbo fôr intransitivo ou, ainda, quando não houver nenhum antecedente, nem expresso nem oculto, será melhor empregar de, porquanto a idéia é sempre de necessidade, de obrigatoriedade. Observemos a correção de Vieira: "... para se conhecerem os amigos, haviam os homens de morrer primeiro e daí a algum tempo ressuscitar". Haviam está aí empregado por tinham, mas, como o segundo verbo é intransitivo (morrer), emprega Vieira, com a meticulosidade de quem muito conhece o idioma, a preposição de em vez do pronome que, ao qual nenhuma função caberia na frase.*" (ALMEIDA, 1969, p.226).

[23] Original: "[…] *em que pese o valor de tão autorizadas opiniões, a verdade é que o levantamento que fizemos [...] não nos permitiria de modo algum negar à expressão ter que + infinitivo o mesmo valor perifrástico reconhecido em ter de + infinitive.*" (ALMEIDA, 1980, p.152)

[24] Original: "[…] *pouco vale argumentar com a função mais comum do que como pronome relativo ou conjunção. Na perífrase de obrigação o que continua a ser um conectivo subordinativo, mas caracteristicamente como preposição, representando mais um deslocamento próprio da língua, de um sistema em movimento, igualmente ao que já ocorreu com durante, mediante, mal, etc. A oposição ter de/ ter que oferece, assim, ao falante da língua, para a expressão modal da obrigação, um recurso bastante expressivo com base na diversidade do ponto de articulação do fonema /d/, oclusiva dental, face a /k/, oclusiva velar.*" (ALMEIDA, 1980, p.152)

(41) *Acabei **de** ler* o livro que o professor recomendou. (I've just read the book that the teacher recommended.)

According to Castilho (2010, p.423, our translation), terminative aspect is a subclass of the imperfective[25] – the qualitative face of the aspectual category – which "marks the final moments of a duration, which is only possible in the periphrases of *acabar de/por,[26] cessar de, deixar de, terminar de* + infinitive"[27]. It can be inferred from the words of Castilho (2010) that the construction [ACABAR + DE + INFINITIVE] is multifunctional, insofar as it denotes a temporal idea of the recent past and an aspectual notion related to the final stage of an event.[28, 29] This overlapping of categories highlights the complexity of the functional domain of the TAM category, as proposed by Givón (1984, 1995), which often results in the impossibility of establishing clear and precise boundaries among time, aspect, and mood, in the different languages, including those belonging to the Romanesque stem.

When studying the coded terminative aspect using auxiliary verbs in Brazilian Portuguese, Bertucci (2010) correlated the combination restrictions of some terminative auxiliaries, namely *acabar, parar,* and *terminar*, with the action classes (Cf. VENDLER, 1967) of the verb inflected in the infinitive. As he noted, these auxiliary verbs are only combined with event verbs: *parar* combines with activity and accomplishments verbs, while *terminar* forms grammatical constructions with accomplishment verbs only. As he speculated, "[…] this greater restriction of *terminar* in relation to *parar* may be happening because *terminar* carries the idea of an end, of the ending of a process and *parar*, the idea of interruption, which may or may not be the end, so it can

[25] However, this is not the view of Travaglia (1985), who justifies his choice, arguing that it is about different levels of representation of the event, given that the terminative notion refers to one of its (final) phases, while the imperfectivity notion concerns its (non)complementation.

[26] As already discussed, we are assuming, like Barroso (1994), that the construction [*acabar + por* + infinitive] denotes not stage aspect (terminative), as proposed by Castilho (*op. cit.*) and Travaglia (1985), but rather placement aspect, more specifically, related to the ordering of actions on the internal time axis.

[27] Original: "*assinala os momentos finais de uma duração, o que só é possível em perífrases de acabar de/por[#], cessar de, deixar de, terminar de + infinitive.*" (CASTILHO, 2010, p.423)

[28] In the understanding of Travaglia (2004, p.37-38, our translation), "[…] the verb *acabar* (+ of + infinitive) expresses recent past tense and, a little by implication, also expresses finished aspect, but not as a marker already specialized in this value. [...] Still in reference to specialization, we observed that the verb *acabar* would have several competitors in its function to mark the finished aspect (*terminar, cessar de, parar de, deixar de, findar, finalizar, largar de*). In this study, we systematically examined two of them: *deixar* and *terminar* and we found that "*terminar*" is losing strength in the paradigm, so much so that it occurs with a very low frequency in the *corpus* (3.52% of the occurrences of the verbs under analysis), while "*deixar*" seems to be acting in the indication of other values and is in a less advanced stage of grammaticalization."

[29] The construction [ACABAR + DE + INFINITIVE] is also multifunctional in Spanish and has some formal restrictions, depending on whether it is temporal or aspectual. According to Bosque and Demonte (1999, p.3334, our translation), "even so, <*acabar de + infinitivo*> with recent past 'puctual' meaning behaves distinctly from what it does with the 'terminative' meaning [...] since only in the second case it fits the elimination of the infinitive and, besides, only in the first case it is possible the repetition of the verb *acabar* in the same construction, and the combination of *acabar* (auxiliar) with *terminar* (auxiliated)."

operate on activities."[30] (BERTUCCI, 2010, p.50, our translation). Regarding the V1 of our example (41), the description propounded by Bertucchi (2010) demonstrated that, as a temporal auxiliary, *acabar* combines with all types of events (activity, achievements, and accomplishments), but, in an aspectual reading, behaves like the verb *terminar*, combining exclusively with verbs of accomplishment, which is predictable, considering the original semantic value of these verb forms and the nature of the verbs of accomplishment. Given the similarity of behavior between the auxiliary verbs *terminar* and *acabar* and the specificity of *parar*, which can be combined with activity verbs, we understand that such auxiliary verbs do not belong to the same class and, therefore, the constructions in which they participate also encode different notions. Therefore, we propose that constructions formed with the auxiliary verb *parar* and with others that are semantically related, such as *deixar* and *largar,* do not encode terminative aspect itself, but rather translate a semantic notion of cessation that, as Bertucci (2010) envisioned, anchored in the idea of interruption expressed by its notional content, denotes a "[…] mixture of time and aspect: insofar as it establishes a contrast between **yesterday** (or **earlier**) and **now** (the moment of speaking), it is temporal; insofar as it indicates that the situation is **finished**, it is aspectual."[31] (TRAVAGLIA, 1985, p.69, our translation), as can be seen in the following contexts:

(42) Com a chegada do inverno, *parei **de** correr* às 6h. (With the arrival of winter, I stopped running at 6:00 AM.)

(43) "[...] o *deixei **de** fazer* por me sobrevirem dores de cabeça com que não podia escrever sem grande moléstia" (I did not do it as I was overcome with headaches with which I could not write without great effort.) (17th century).

(44) Cansado de ser reprovado em entrevistas, *larguei **de** procurar* emprego e decidi abrir meu próprio negócio. (Tired of getting turned down in interviews, I stopped looking for a job and decided to start my own business.)

It should be noted that, contrary to what occurs in modal auxiliary verb constructions, when it comes to coding aspectual notions, the preposition DE only occurs in auxiliary constructions that still preserve their original semantic value and where the expression of terminative aspect is restricted to those auxiliary verbs that carry the notion of ending or finishing something in their lexical content, which vetoes its concomitance with activity verbs.

Considering the complexity of the description, it seems to us not only legitimate, but also opportune, especially from the didactic point of view, to systematize the

---

[30] Original: "[…] *essa maior restrição de terminar em relação a parar pode estar acontecendo porque terminar carrega a ideia de fim, de término propriamente dito de algum processo e parar, uma ideia de interrupção, que pode ser o fim ou não, por isso pode operar sobre atividades.*" (BERTUCCI, 2010, p.50)
Activity verbs denote events that last for a certain time, but which, unlike accomplishment verbs, do not require an end.

[31] Original: "[…] *mistura de tempo e aspecto: na medida em que estabelece um contraste entre **ontem (ou antes)** e **agora** (momento da enunciação) é temporal; na medida em que indica que a situação é **acabada** é aspectual.*" (TRAVAGLIA, 1985, p.69).

generalizations reached here about our second inquiry before we proceed with the analysis of the third question we formulated: Is there specificity of function related to the type of prepositional link or do the same prepositions participate in the coding of different functionalities? Our analysis showed that, although the same preposition can participate in the coding of more than one functionality, even having more than one preposition participating in the coding of the same function, there are specializations related to each one of them, as summarized in the table below:

Table 1 – Correlation between prepositional link and
functionality in Portuguese auxiliary verb constructions

Preposition	Functionality whose coding the preposition participates
A	• *Stage aspect* (inceptive): combined with auxiliary verbs that preserve their original meaning: *começar **a** chorar; principiar **a** correr...* • *Stage aspect* (inceptive) + *extension aspect* (duration and iteration): combined with auxiliary verbs semantically emptied of their original meaning: *pegou **a** chorar; disparou **a** gritar; danou **a** mentir....* • *Extension aspect* (durative): productive in European Portuguese but not in Brazilian, whose preference is for the nominal form of the gerund: *fiquei **a** pensar* (= *fiquei pensando*) It has greater productivity in the second function.
DE	• *Stage aspect* (terminative): combined with auxiliary verbs that preserve their original meaning and accomplishment verbs: *terminar **de** ler o livro; acabar **de** comprar o carro...* • *Semantic notion of cessation*: combined with auxiliary verbs that preserve their original meaning and with activity verbs: *parar **de** nadar; deixar **de** caminhar...* • *Stage aspect* (inceptive) + extension aspect (duration and iteration): combined with auxiliary verb *dar* determinate subject, a context in which it competes with PARA: *a criança deu **de/para** ter* medo de escuro... • Modality: combined with semantically emptied auxiliary verbs, a context in which it competes with QUE, in the expression of deontic modality: *hei **de** conseguir* (desejo); *tenho **de/que** conseguir* (obligation) It has greater productivity in the coding of modality.
PARA	• *Stage aspect* (inceptive) + extension aspect (duration and iteration): combined with the auxiliary verb *dar* with determinate subject*: a criança deu **para** ter* medo de escuro... • *Epistemic modality:* combined with the auxiliary verb *dar* with indeterminate subject: *não deu **para** ler* o livro durante o final de semana. • *Temporality* (imminent future): in the construction with the auxiliary verb *estar*: *a reforma da previdência está **para** ser votada*. Combines with two auxiliary verbs only: *dar* and *estar*, the first one being more semantically emptied.
POR	• *Placement aspect* (order)**:** combined with auxiliary verbs that preserve their original meaning and with accomplishment or activity verbs: *acabou **por** comprar* um carro usado; *terminou **por** andar* sem destino até o amanhecer... • *Extension aspect* (durative): in the construction with auxiliary verbs *estar* and *ficar*: *sua lição ainda está/ficou **por** fazer*, meu filho!

Source: Author's elaboration.

Restriction of preposition elimination in the context of auxiliary verb constructions

If, in contemporaneity, the elimination of prepositional links in contexts of auxiliary verb constructions is more limited to situations of orality in the archaic period of the history of the Portuguese language, which encompasses, in the chronology of Mattos & Silva (2001), the interval of time between the thirteenth and fifteenth centuries, it is also possible to document it in written texts, as illustrated by these excerpts taken, respectively, from the work, *Orto do Esposo* (end of the fourteenth century to the beginning of the fifteenth century), cited by Almeida (1980), and from the work *Vidas de Santos de um Manuscrito Alcobacence* (1200-1300), an integral part of the databank of the *Portuguese Corpus* (https://www.corpusdoportugues.org/):

(45) "[...] e este seu filho, pero era filho de bõo padre, *começou husar* mal de ssy em guysa que se queria tornar a adorar os ydollos" ("[..] and this, your son, but was the son of the good priest, began to use himself poorly as he wanted to begin to adore idols".) (*Orto do Esposo,* 183 *apud* ALMEIDA, 1980, p.49).

(46) "O qual logo abrio sua boca e per espiritu sancto *começou ffallar* pallavras consolatórias [...]" ("Who soon opened his mouth and by the holy ghost began to speak words of consolation [...]".) (*Vidas de Santos de um Manuscrito Alcobacence, Corpus do Português*)

The absence of the preposition in the constructions highlighted in (45) and (46) does not clearly illustrate a specific regency of the period, but rather the elimination that we are discussing in this paper, given that, in the same works, there is the record of the construction with the indirect incidence of the preposition A, along the lines of the contemporary written use, as well as the preposition DE, the most productive in that period:

(47) "E, estando ele deitado em oraçõ em aquelle loguar, *começou* sobitamente ***a dormir***" ("And, as he was lying in prayer in that place, he subtly began to fall asleep".) (*Orto do Esposo,* 167 *apud* ALMEIDA, 1980, p.49).

(48) "E *começarõ **de** andar* per hũu valle que era muy escuro" ("And they began to walk through the very dark valley.) (*Vidas de Santos de um Manuscrito Alcobacence, Corpus do Português)*.

Faced with this data, the question at hand is how to explain the fact of the inceptive aspectual construction being formed with both the presence and the absence of the preposition. Moreover, considering that, during the period, there was a certain competition between the prepositions A and DE in the context of the construction, as affirmed by the data presented in (47) and (48), it is impossible to know, precisely, if both could be deleted or if only the preposition A, as occurs in contemporary constructions,

could be deleted.[32] According to the teachings from Maurer Júnior (1959), the fact that the preposition governs the infinitive constitutes the greatest innovation of the syntax of the vulgar verb, opening the door to innumerous new applications of this nominal form.

> Thus, <u>since the archaic era, the vulgar language distinguished itself from that of the aristocratic and literary, which admitted the infinitive</u> as a complement of nouns and adjectives, or together with verbs, to express, through the simple juxtaposition, a wide range of relationships that in the common nouns was indicated by the causal reflection. […] The writers of the era, as they could not escape from the tendency of the spoken language, avoided, nevertheless, a grave solecism, suppressing the preposition that governed the infinitive in the speech of the people. (MAURER JÚNIOR, 1959, p.185-186, our translation, our highlight)[33].

This citation shows that the use of the preposition in contexts of auxiliary verb constructions, such as those studied herein, was initially considered to be a vulgar expression. In oral speech, which was refuted in the writings of the authors of the day. This information can even be evoked to explain the elimination of the preposition in contexts (45) and (46) presented above, but it clearly does not apply to the contemporary eliminations, as shown below:

(49) a. A criança *começou **a** chorar* de fome. (The child began to cry of hunger.)
      b. A criança *começou chorar* de fome.
(50) a. A criança *danou **a** chorar* de fome. (The child began to cry wildly of hunger.)
      b. A criança *danou chorar* de fome.
(51) a. A criança *tornou **a** fazer* birra, para desafiar a mãe. (The child began to pout again, to challenge his mother.)
      b. A criança *tornou fazer* birra, para desafiar a mãe.

---

[32] As regards the competition between the prepositions A and DE in this context, we believe that both prepositions may have been eliminated during the period, given these two pieces of data taken from the work *Vidas de Santos de um Manuscrito Alcobacence* (*Portuguese Corpus),* in which, though dealing with the same syntactic environment, either the phrase was created by direct incidence or by mediation of the preposition DE: (i) "E disy *começarõ andar* per hũa carreira torta muy maa pella qual a alma nom avya outro lume" ((i) And from this they began to travel down a very winding and nasty road, down which the soul had no other light") e (ii) "E *começarõ **de** andar* per hũu valle que era muy escuro". ((ii) And they began to walk through a very dark valley.)

[33] Original: "*Assim, <u>desde a época arcaica a língua vulgar se distinguia da aristocrática e literária em que admitia infinito</u> como complemento de nomes e adjetivos, ou <u>junto a verbos, para exprimir, pela simples justaposição, diversas relações que nos substantivos comuns eram indicadas pela flexão casual. [...] Os escritores da época, não podendo escapar à tendência da língua falada, evitam, contudo, um grave solecismo, suprimindo a preposição que regia o infinito no falar do povo.*" (MAURER JÚNIOR, 1959, p.185-186, grifo nosso).

(52) a. A criança *deu **para** fazer* birra para chamar a atenção da mãe. (The child started making a scene, pouting to call his mother´s attention.)

b. A criança *deu **p'ra** fazer* birra p(a)ra chamar a atenção da mãe.

c. A criança *deu **pa'** fazer* birra p(ar)a chamar a atenção da mãe.

d. * A criança *deu fazer* birra p(ar)a chamar a atenção da mãe.

(53) a. A criança *tem **de** comer* frutas e legumes. (The child has to eat fruits and vegetables.)

b. A criança *tem **que** comer* frutas e legumes.

c. * A criança *tem comer* frutas e legumes.

(54) a. A criança *terminou **de** fazer* a lição e foi brincar. (The child stopped doing his homework and went out to play.)

b. * A criança *terminou fazer* a lição e foi brincar.

(55) a. O orador *começou **por** agradecer* a presença do público. (The speaker began by thanking the public for their presence.)

b. *O orador *começou agradecer* a presença do público[34].

(56) a. Todo o trabalho ainda *está **por** terminar*. (All of the work is still to be finished.)

b. * Todo o trabalho ainda *está terminar*.

The analysis of the contexts of (49) and (56) suggests that a factor capable of allowing the elimination of the prepositional link in some contemporary constructions is the type of preposition that participated in the composition, that is, only the preposition A can be eliminated from its context without harming the proper formation of the sentence. Consequently, as discussed in the previous section, this also restricts the optionality of the preposition to the functionalities of whose coding of the preposition "A" participates, that is, stage aspect (inceptive) and extension aspect. Once the preposition, which allows for elimination, is identified, it is up to us to understand why only it allows for the elimination. What appears is the complex phenomenon upon which acts a confluence of factors involving phonological, prosodic, morphological, and semantic questions, as we will explore below.

First, it is important to observe that, the more grammaticalized the construction, the greater the cohesion among its constituent elements and, in this case, the preposition is so integrated to the auxiliary verb that it begins to behave as if it were part of it, along the lines of a morpheme, subsequently transmuting the meaning. No Portuguese speaker would hesitate to admit that *ter* and *ter de/que* ("to have" in English), for example, are two distinct verbal forms that denote different meanings in the language. The same can be said of *começar* and *começar por* (*to begin*), de *acabar* e de *acabar de* (*to end*)... It is as if these prepositions, upon being juxtaposed with the auxiliary verb, then became part of it, behaving as a morpheme with a grammatical meaning from the semantic

---

[34] The proper formation of this construction is only assured by an inceptive reading, but not by the ordination that the presence of the preposition POR encodes.

categories of aspect, mood, and time, which, to a certain extent, can be interpreted as an extension of the old causal role that it played in Latin. If our interpretation of the fact is correct, upon integrating the auxiliary verb constructions, the preposition would be advancing in its *continuum* of grammaticalization, much like Lehmann (1995) conceived: synthetization > morphologization > phonological reduction > stage zero. Its apposition to the auxiliary verb to mark grammatical categories would correspond to, in this process, the morphologization stage. In this stage, the formation results from a single phonological word,[35] here understood in the meaning proposed by Câmara Júnior (1975, p.38, our translation): "prosodic entity, characterized by an accent and two degrees of possible tones before and after the accent." The fact is that, when affixed to the auxiliary verb, forming with it a phonological word, the preposition A becomes the final unstressed vowel of the verbal form and, consequently, it weakens, thus becoming subject to the same morphophonological processes[36] that this vowel undergoes in an environment of an intervocabulary junction.

In this environment, a context that is not only favorable but also categorical for the elimination of the preposition is, for instance, that in which the verb *começar (begin)* appears conjugated in the third person singular: *começa a dizer (begin to say)*. In this case, the crasis of the final unstressed vowel of the auxiliary verb with the preposition, resulting in a construction of direct incidence auxiliary verb, along the lines of those recorded in archaic Portuguese: *começa dizer (begin to say)*. This environment of the vocabulary juncture can become even more favorable to the elimination, if the nominal form of the infinitive also begins with the vowel /a/, as seen in *adorer (to adore)*, for instance (*começa a adorer (begin to adore)*), which, by means of an external Sandhi, will result, through the concomitance of the double crasis, in a single phonological term (/komɛsado'ra/). This, however, is not the most productive environment for the elimination of the preposition in the constructions described herein, which tend to be employed, for the most part, with the auxiliary verb conjugated in the past simple, whose conjugations of the grammatical person and verbal time result in a form that ends in the falling diphthong: *começou **a** estudar, comecei **a** estudar, começaram **a** estudar (He/She began to study, I began to study, they began to study).* Given that this no longer refers to a context in which a crasis is used, how can one explain the similar elimination of the preposition A? For Câmara Jr. (1998, p.39, our translation), "[…] the occurrence

---

[35] In the understanding of Câmara Jr. (1998), which we endorse here, the dependent forms, such as the prepositions and the particle *que*, concurrent of the preposition in our modal constructions, are morphic words, "[…] but they do not constitute in themselves phonological words. To the contrary, they create a phonological word with a free form that follows or precedes it." (CÂMARA JUNIOR, (1998, p.37, our translation).

[36] According to warnings by Cagliari (1997, p.62, our translation), "[…] by describing some contexts, there is the need to take into account […] not merely the preceding and subsequent words, but rather the fact of the context being or not linked to the external limits of the word (also called the *intervocabulary junction*), or belong to a specific lexical or syntactic category (for example, verb in the infinitive, noun, etc.).", as is the case of the constructions studied herein. When this happens, "[…] one has, as the starting point for considerations of this nature, not the phonetic, but the morphological facts, such as the *basic form* of the morphemes. When a basic lexical form serves as the motivation for a phonological rule, a *morphophonological process* occurs." (CAGLIARI, 1997, p.62-63, our translation).

of the variation in inflectional in the first element of speech[37] is then distinguished, effectively, from a formal unitary word, because, in Portuguese, the formal word can only be a variation in inflection in its final part."[38] Given the circumstances, we will have to analyze the preposition in the context of its group of strength, defined in this manner due to its intensity in the emission of its syllables, that is the accent. In this context, one should initially consider two types of syllables: (i) the *stressed* (syllable of exceptional strength) and (ii) the *unstressed* (less stressed syllable). Since the unstressed syllables present a wide range of emission deficiencies, they must be subcategorized according to this deficiency: (i) *subtonic* (unstressed syllable of a derived word, but that was stressed in the word's primitive form), (ii) *pretonic* (preceding the stressed syllable), and (iii) *post-tonic* (following the stressed syllable) and/or *final unstressed syllable*. To build an accent or prosodic agenda of the words, Câmara Jr. (1998) proposes to attribute a distinct numeral to each one of these syllable types, with 3 corresponding to the greatest intensity of the scale and reserved, therefore, to the stressed syllable. In this gradation deficiency, the subtonic would be marked by the numeral 2, the pretonics by the numeral 1, and the post-tonics and endings would be equally marked by the numeral 0. According to Câmara Jr. (1998, p.35), "[…] in the strength group, only the stressed syllable of the final phonological word maintains the maximum accent 3. That of each of the preceding words remains with a more stressed accent."[39], as illustrated below, through one of our constructions:

$$[\text{CO ME ÇOU } \mathbf{A} \text{ CHO RAR}]$$
$$1 \quad 1 \quad 2 \quad 0 \quad 1 \quad 3$$
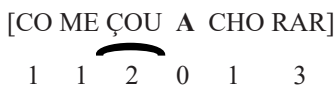
The analysis of this accentual guideline of the construction above is elucidative enough to demonstrate that the preposition A is the only element of the group devoid of prosodic strength,[40] which, in itself, already makes it a potential segment for elimination. Add to this the fact that it links, from the prosodic point of view, to an auxiliary form that preserves its original semantic content, in addition to the fact that it is the smallest of all prepositions of the Portuguese language, composed of a single phoneme and, as predicted by the parameter of integrity (Cf. LEHMANN, 1995), the substantial size of a sign is strictly connected to its semantic and phonological matrix. In this sense, of all

---

[37] Câmara Júnior (1998, p.38, our translation) thus distinguishes the *juxtaposition* of the *locution*: "Along with the concept of juxtaposition, which is a formal word composed of two phonological words, we have the concept of "locution", so that they are strictly two formal words." which create a single phonological word.

[38] Original: "[…] *a ocorrência de variação flexional no primeiro elemento da locução logo a distingue, com efeito, de um vocábulo formal unitário, porque em português o vocábulo formal só pode ter variação flexional na sua parte final.*" (CÂMARA JUNIOR., 1998, p.39).

[39] Original: "[…] *no grupo de força, só a sílaba tônica do último vocábulo fonológico mantém o acento máximo 3. A de cada um dos vocábulos precedentes fica com acento mais atenuado.*" (CÂMARA JUNIOR, 1998, p.35).

[40] In theory, this is predictable, considering that the preposition is a stressed particle that is clitic to the auxiliary, with which it forms a phonological word, and which "the initial and final phonemes within a strength group remain poorly marked as such" (CÂMARA JUNIOR, 1998, p.35, our translation).

of the Portuguese prepositions, A is the most semantically empty and, by itself, evokes no type of semantic relationship, contrary to that which occurs with DE, for example, which is highly polysemic and capable of remitting to the relationships of possession, of origin, of place, with "PRA, which remits to the idea of direction; or even POR, which evokes the idea of route or displacement. This may well be why Pontes (1992, p.24, our translation) considers this a context of use of the preposition A, whose "meaning is not easy to precisely define"[41]. In addition to the fact that this preposition is devoid of prosodic strength and semantic matrix, it is clitic to a rising diphthong [ow], which constitutes, in the words of Câmara Jr. (2011, p.211, our translation) a categorical context of the monophthongization[42]:

> […] phonetic change that consists of the passage of a diphthong (v.) to a simple vowel, such as the passage in Latin of the ae to /è/ and in vulgar Latin from *au* to *o* (*pauper* > **poper;* cf. port. *pobre*). To highlight the phenomenon of monophthongization, what is often called MONOPHTHONG is the simple resulting vowel, primarily when the spelling continues to indicate the diphthong, and it is still carried out in a more careful language. Among us, there is, in this sense, the monophthong *ou* /ô/, in any case, and then /a/, ei /ê/ before a 'sizzling' consonant sound; exs.: *(p)ouca, (b)oca, (c)caixa,* como *acha, (d)deixa),* as in *fecha.*[43]

In this sense, the enclitics of the preposition to the monophthong auxiliary [kome'so] promotes the creation of a new vowel group, which can be articulated both as a hiatus and a diphthong. The fact is that the two articulatory possibilities become dissonant, because:

> […] the succession of the two syllable vowels, contrary to the so-called "syllabic norm", which consists of the regular and undefined alternation

---

[41] Original: "*significado não é fácil precisar*" (PONTES, 1992, p.24).

[42] Paul Teyssier (1997, p.63, our translation) also attests to how traditional and productive the monophthongization of [ow] in [o] is in the history of the Portuguese language. According to his report, "this monophthongization most likely began to appear in the eighteenth century. It invaded the entire South and most of the Central regions of Portugal, but in the rest of the country, that is, once again, in the North, the old diphthong *ou* [ow] continues alive." "Portuguese became the language of Brazil in the mid-sixteenth century, in other words, at a date in which the first evolutions […] had already occurred: the elimination of numerous vowel connections […], unification of the singular form of the words like *mão, cão, leão* [...], the maintenance of the distinction between /b/ and /v/ [...], the simplification of the syllable systems […]. In all of these points, the *coined* Brazilian generalized the Portuguese norm of the Center-South, in which the particularities of the North were eliminated. And, during part of the colonial period, it continued to evolve according to that derived from the European Portuguese: the transformation of the monophthong *ou* in [o] [...]" (TEYSSIER, 1997, p.99, our translation).

[43] Original: "[…] *mudança fonética que consiste na passagem de um ditongo (v.) a uma vogal simples, como a passagem em latim de ae para /è/ e em latim vulgar de au para o (pauper > *poper; cf. port. pobre). Para pôr em relevo o fenômeno da monotongação chama-se, muitas vezes, MONOTONGO à vogal simples resultante, principalmente quando a grafia continua a indicar o ditongo e ele ainda se realiza numa linguagem mais cuidadosa. Entre nós, há nesse sentido o monotongo ou /ô/, em qualquer caso, e ai /a/, ei /ê/ diante de uma consoante chiante; exs.: (p)ouca, (b) oca, (c)caixa, como acha, (d)deixa), como fecha.*" (CÂMARA JUNIOR, 2011, p.211)

of a phoneme from the small opening to a larger opening […]. From this stems the tendency, which can be said, in general, to suppress the vowel meetings in hiatus, in one or another form. One of the mediums used is the evaluation of the existing differences, whose result is a diphthong. (BASSETO, 2010, p.44, our translation).[44]

This option, however, is even less reasonable in the case of our construction, since the vowel sequence /o/ + /a/ produces the odd diphthong in Portuguese whose glides are, respectively, /y/ and /w/. In this sense, so as not to infringe upon the structural tendency, since "against the environmental strength, there is, in languages, the specific structural strength of each system"[45] (CAGLIARI, 1997, p.15, our translation), it maintains, in the syllabic enunciation, the monophthong with the subsequent fall of the final unstressed segment represented by the preposition A. That happens because in the game of expiatory strength, the preposition is eliminated by occupying a syllabic position. This option preserved, to a certain extent, the linguistic tradition, considering that, since Latin, a tendency toward the reduction of the diphthong has been observed.[46] In addition, the vowel syncope preserved the phonetic law.

> According to that verified by neogrammarians, the stressed vowel, both in Latin and in Romance languages, possessed a great stability and rarely underwent major changes. It is called the *law of persistence of the stressed syllable*. The unstressed vowels of the internal syllables, however, can undergo syncope […]. The factors that caused, at times concomitantly, are the extension of the word, the faster or slower speed of the elocution, the nature of the surrounding phonemes, and, even greater, the intensive accent. […] This tendency is accentuated in vulg. Lat. and in Romance languages, since the syncope is a primarily popular and familiar character, stemming from the vulgar variety and, consequently, from the Romance languages (BASSETO, 2010, p.48, our translation).[47]

---

44 Original: "*a sucessão de duas vogais silábicas contraria a chamada "norma silábica", que consiste na alternância regular e indefinida de um fonema de pequena abertura e de outro de grande abertura [...] Daí resulta a tendência, que se pode dizer geral, de suprimir os encontros vocálicos em hiato, de uma ou de outra forma. Um dos meios utilizados é o aprofundamento das diferenças existentes, cujo resultado é um ditongo* (BASSETO, 2010, p.44)"

45 Original: "*contra a força ambiental, há nas línguas a força estrutural, própria de cada sistema*" (CAGLIARI, 1997, p.15)

46 Basseto (2010, p.41, our translation) reports that "[…] the Lat. Diphthongs were inherited from the ind. -eur., with clear tendencies toward reduction. Meillet and Vendryes observed that all of the ind. -eur. diphthongs remained nearly intact at the date of the oldest Lat. Documents, and were, however, simplified in the discourse of the history of the language (*Traité de gram. comp.des lang.clas.,* p.112). The reduced number of diphthongs from the clas. Lat. led Friedrich Diez to the conclusion that this norm fed 'an accentuated dislike of the diphthongs' (*Gram. des lang. rom.,* I, 184)."

47 Original: "*Segundo verificaram os neogramáticos, a vogal tônica, tanto latina como românica, possui grande estabilidade e dificilmente sofre maiores alterações. É a chamada lei da persistência da sílaba tônica. As vogais átonas das sílabas internas, porém, podem sofrer síncope [...]. Os fatores que as causam, por vezes concomitantemente, são a extensão do vocábulo, a maior ou menor rapidez na elocução, a natureza dos fonemas circunvizinhos e, com maior*

In addition to this combination of facts, including diachronic factors, affirming that strength from the past continues acting in the present, the preservation of the monophthong when faced with the elimination of the preposition A is also a means through which to shape the phonological word to the syllabic pattern that is predominant in words from the Portuguese language, the so-called free or open syllables, composed structurally of a consonant and a vowel (CV). The maintenance of the basic syllabic pattern, coupled with the prosodic agenda of the construction and with the semantic load of the prepositional link seem to be, effectively, relevant factors to determine the elimination, especially when we substitute the preposition A for others that are not eliminated from the context:

<div align="center">

[A CA BOU **PO(R)** COM PRAR]
1 1 2   2   1   3

[TEM **DE** FA ZER]
2   1   1 3

[DEU **PA (RA)** BE BER]
2   2   0   1 3

</div>

When we examine the prepositions POR, DE, and PARA, which cannot be eliminated from the context of the construction, we perceive that, what weighs in their favor is the fact that they have a greater strength than the preposition A to conform to the phonological environment in which they occur, to the CV syllabic pattern, in addition to being replete with a broader semantic load than preposition A. It is important to observe that the preposition PARA, which contains the biggest sound load of the group and, consequently, a bigger notional content, only combines with verbs whose original meaning has been lost, which indicates that the semantic matrix of the prepositional item is relevant to the functional coding of the constructions and, therefore, also determines the possibility of it being or not eliminated from the context.

Before we pass to the final subsection, where which we reflect on the factors that determine the selection of the preposition in the context of the construction of the auxiliary verb, it is important to contemplate the implications of our analysis when considering the fact that, as native speakers of Portuguese and Galician have reported, during our presentation at the IV International Congress on Historical Linguistics (Homage to Ivo Castro), which took place in Lisbon from July 17-21, 2017, in these two Romance languages, contrary to that which occurs in Brazil, not even the preposition A can be eliminated from the construction without compromising the grammatical structure of the sentence. Considering our analysis, according to which the (im)possibility of elimination is heavily motivated by prosodic questions, we assess that the fact that the Portuguese from Portugal and the Galician preserved the preposition A in its

---

preponderância, o acento intensivo. [...] Essa tendência se acentuou no lat. vulg. e nas línguas românicas, uma vez que a síncope é um fenômeno de caráter sobretudo popular e familiar, próprio portanto da variedade vulgar e, em consequência, das línguas românicas." (BASSETO, 2010, p.48).

construction can be explained by a phonetic difference of the timbre of the vowel A in the referred languages, when in final unstressed position, which, clearly changed the stress pattern of the combination. It is well-known that "[…] it is primarily in the pronunciation of the vowels that Brazilian Portuguese distances itself, both due to its conservatism as well as due to its innovations, from European Portuguese."[48] (TEYSSIER, 1997, p.104, our translation). Moreover, such a difference can be affirmed in the unstressed positions, due to the fact that the Portuguese stress is "[…] intensive, but not violent. It is much stronger in Portugal then in Brazil, with a broad contrast between the stressed and the unstressed syllable, which in Brazil cannot be seen."[49] (CÂMARA JUNIOR, 1976, p.33, our translation). In the specific case of the final unstressed [a], "although shorter than in the stressed position, remains very open"[50] (TEYSSIER, 1997, p.100, our translation) in Brazil, in addition to it here having neutralized the opposition of the open and closed timbre of this phoneme in the pretonic syllable, which is also a possible context for the preposition A in auxiliary verb constructions (*começou **a** dar (began to give)*). This would be, for example, a context in which the Lusitanian articulation would not allow the elimination of the preposition in the manner in which it occurred in Brazil.

Determining factors for the selection of the preposition in the context of auxiliary verb constructions

A final question about which we proposed to reflect in the context of the auxiliary verb constructions of indirect incidence concerns the selection of the preposition in a given syntactic environment. If we consider that, in vulgar Latin and, consequently, in the Romance languages, the role of the regimen originally fit within the nominal form of the infinitive, given that, in constructions such as *quero trabalhar* (Portuguese), *quiero decir* (Spanish), *il doit partir* (French), and *passiamo capire* (Italian), "we have an infinitive that serves as a direct complement to another verb"[51] (MAURER JUNIOR, 1959, p.183, our translation), we tend to believe that it is the role of the auxiliary, as a remainder of its old lexical function, to select the nominal form that will compose the construction with it (COELHO, 2013). From this, we contemplated whether or not the preposition would also be selected by the auxiliary form, a hypothesis that initially led us to assume that the construction of the auxiliary verb of direct incidence would be restricted to auxiliary forms that were originally direct transitive auxiliaries, given that those of indirect incidence would have been grammaticalized in order to aid indirect transitive verb forms. Nevertheless, the processes of linguistic change are

---

[48]   Original: "[…] *é principalmente na pronúncia das vogais que o português do Brasil se distancia, tanto pelo seu conservadorismo como pelas suas inovações, do português europeu.*" (TEYSSIER, 1997, p.104).

[49]   Original: "[…] *intensivo, mas não violento. É muito mais forte em Portugal do que no Brasil, com um grande contraste entre sílaba tônica e sílaba átona, que no Brasil não se verifica.*" (CÂMARA JUNIOR, 1976, p.33).

[50]   Original: "*embora mais breve que em posição tônica, permanece muito aberto*" (TEYSSIER, 1997, p.100).

[51]   Original: "*temos um infinito que serve de complemento direto a outro verbo*" (MAURER JUNIOR, 1959, p.183).

overly complex to be limited to this overgeneralization and prodigal are the examples of constructions of auxiliary verbs of indirect incidence with related auxiliary verbs, direct transitives, bitransitives, or even unaccusatives, as illustrated, respectively, in the data below:

(57) a. O trabalho *ficou **por** terminar*. (The work went unfinished)
      b. O bebê *está **para** nascer*. (The baby is about to be born)

(58) a*. Hei **de** conseguir* o emprego. (I will get a job.)
      b. *Tenho **de** terminar* o trabalho. (I have to work)
      c. Os moradores *pegaram **a** reclamar* das atitudes do novo porteiro. (The residents began to complain about the attitudes of the new doorman.)

(59) Aquela criatura, antes tão meiga, *deu **para** ser* malcriada com qualquer um que lhe dirija a palavra. (That so-and-so, who used to be humble, began being so rude with any who spoke to him.)

(60) a. A criança *começou **a** chorar* de repente. (The child suddenly began to cry.)
      b. *Acabamos **de** ouvir* a notícia no jornal. (We've just heard the news on the news program.)

We cannot lose sight, however, that the grammaticalization process from which the auxiliary verb constructions result involve reanalysis and that, therefore, the key to understanding the presence of the preposition in the contexts illustrated above can only be identified, if they can be, historically, that is, trying to recover the syntactic context that promoted the reanalysis. To the best of our knowledge, the auxiliary verb constructions of indirect incidence have a privileged locus for their emergence within contexts of purpose. Running in line with our data is the information from Maurer Júnior (1959, p.184, our translation), according to which

> […] the idea of the end is associated with the primitive form of the infinitive, if the commonly admitted interpretation that it deals with an old dative (*amari*) or locative (*amare*) is correct. Bennett admits that the idea of *direction* and *end* have gone from the dative to the locative.[52]

Our data more specifically point out that the process of reanalysis occurs in a context of purpose, often favored by the anteposition of complements selected by the nominal form of the infinitive, as shown below, through the analysis of these data extracted from the *Portuguese Corpus* (https://www.corpusdoportugues.org/):

(61) "[...] nẽhũra cousa a mĩ *ficou por dar*" ("Nothing was left to be given to me") (13th century, Notary Texts)

---

[52] Original: "[…] *a idéia de fim está associada à forma primitiva do infinitivo, se a interpretação comumente admitida de que se trata de um velho dativo (amari) ou locativo (amare) é correta. Bennett admite que a idéia de direção e de fim tenha passado do dativo para o locativo.*" (MAURER JÚNIOR, 1959, p.184).

(62) "[...] nẽhũa cousa pode séér dicta perfecta nẽ acabada quando algũa cousa dela *fica pera fazer*" ("Nothing can be considered perfect or finished when something of it is left undone".) (14th century, Documents from the Santa Cruz de Coimbra Monastery)

(63) "[...] e muitos clerigos e leygos que hy *estavã pera lhe fazerẽ* honrra asi como fazẽ a homẽ morto [...]" ("[...] and many priests and laymen who were there to pay homage to him like they do to a dead man […]") (18th century, Lives of the Saints from an Alcobaça Manuscript)

(64) "Este foi todo o fundamento das honras que me fez e licenças que me *deu para pregar e baptizar* por todo o reino" ("This was the entire basis of the honor that was given to me and the permission that it bestowed upon me to preach and baptize throughout the entire kingdom".) (17th century, History of the Life of the Priest, St. Francis Xavier)

(65) "[...] posto que as ocasiões de serviço, ou chamado serviço, de V. M. têm sido tantas estes dois anos que não tiveram os pobres índios lugar de lograrem os seis meses que V. M. lhes *manda dar para acudirem* a suas lavouras e casas e para conhecerem que não são cativos [...]" ("[...] since the occasions of service, or the so-called service, of Your Majesty has been so great these last two years that the poor Indians were not able to take advantage of the six months that Your Majesty grants them to take care of their harvests and homes and to understand that they are not captives […]") (17th century, letters from the Priest, Antônio Vieira).

(66) "Amigos e vassalos, bem sabees em como me enlegestes por vosso rey e o juramento que me *destes de fazer* dereito e justiça assy aos grandes como aos pequenos" ("Friends and vassals, you well know how you elected me as your king and the oath I have sworn to do justice to both the powerful and the weak".) (14th century, General Chronical of Spain Espanha, 1344).

Despite the difference both in the transitivity of the reanalyzed verbal forms as future auxiliaries, as well as the prepositions that precede the nominal form of the infinitive in the data presented in (61) to (66), we observed the verb FICAR, followed by the prepositions POR and PARA and of a final infinitive whose respective complements appear antepositioned. Despite the ambiguity of the forms, which is common in the grammaticalization and reanalysis processes, it is already possible to predict the durative notion that the construction [FICAR + POR/PARA + INFINITIVE] began to denote. In (63), the verb ESTAR is still not reanalyzed as the auxiliary, but its circumstantial complement is antepositioned, as is the dative complement of the infinitive FAZER, which favors the juxtaposition of the forms for a future reanalysis. Just the same, it is possible, in the relation of purpose introduced here, to identify the notion of imminence that the construction [ESTAR + PARA + INFINITIVE] brought to contemporaneity, given that the action of purpose (*give homage to the deceased*) has still not been accomplished, although all were there to achieve exactly this. In the passages in (64)

and (66), we identify occurrences of the verb DAR, today grammaticalized as an auxiliary in modal and aspectual constructions, as well as in contexts of purpose with an antepositioned dative complement. To the contrary of the other syntactic environments, in this specific case, it is still not possible to precisely define if the latent functionality is modal or aspectual, mainly because, as we have demonstrated in the case of the construction [DAR + PARA/DE+ INFINITIVO], the grammatical function results both from the conjunction as well as from the type of subject selected and its semantic role, which should necessarily be of agent, as it deals with the connotation of aspect.

In addition to the context of purpose analyzed above and already discussed in the previous section, which explains the presence of the preposition in modal constructions with the auxiliary verbs TER and HAVER as the remainder of a nominal complement whose regent noun was eliminated after its illocutionary content was incorporated within the construct, our study identified a relative productivity of unaccusative verbs that were grammaticalized as auxiliary verbs in aspectual constructions. In the specific case of such verbs, in addition to the archaic data that we have already retrieved, in most cases, the already grammaticalized construction, it does not seem worthwhile to assess the reanalysis of a context of purpose introduced by a preposition plus infinitive, as shown through the diachronic data below, collected from the *Portuguese Corpus* (https://www.corpusdoportugues.org/):

(67) "[...] adorey o lenho da cruz e *começey a andar* meu caminho [...]" ("[...] I adored the wood of the cross, and I began to follow my path [...]") (13th century, Lives of Saints from an Alcobaça Manuscript)

(68) "E esta moeda era de tantos dinheiros o maravidil que *chegava a valer* o maravidil tanto como hũru maravidil d'ouro [...]" ("And this coin, the *maradivil*, was so valuable that it started to be as valuable as gold [...]") (14th century, General Chronical of Spain from 1344)

(69) "E, despois que esto *acabou de dizer*, levantousse o Cide e foilhe beyjar a mão" ("And after he finished doing this, Cide stood up and went to kiss his hand") (14th century, General Chronical of Spain from 1344)

(70) "Voltando para casa mui satisfeita, ordena aos criados que, assim ensopados, sirvam à mesa, onde o marido *acabou por conhecer* a extravagante bizarria da sua querida e formosa esposa" ("Returning home quite satisfied, she orders her soaked servants to set and serve the table, which led her husband to witness the eccentricity of his dear and beautiful wife") (18th century, Letters from Cavaleiro de Oliveira).

(71) "E *terminou por confessar* que desde os nove anos não voltara a Sintra" ("And he ended up confessing that he had not returned to Sintra since he was nine") (18th century, Eça de Queiroz).

(72) "Quando Michelangelo *terminou de esculpir* a encomenda feita por um cardeal francês, a Pietà, o mundo inteiro descobriu que o artista possuía um raríssimo dom: o germe da beleza" ("When Michelangelo finished sculpting

the request made by the French Cardinal, Pietà, the entire world discovered that the artist had a very rare gift: the seed of beauty") (20th century, The war of imaginations).

The excerpts (67) to (72) are composed of constructions with unaccusative verbs grammaticalized in the auxiliary function. All of these show aspect, be it of stage (Cf. (67), (69), and (72)), be it of collocation (Cf. (68), (70), and (71)). In some of these, such as in (69), for example, it is possible to identify how in those whose process of reanalysis involves a context of purpose, the anteposition of the complement of the verb form of the infinitive. In (68), we identified the argument of the verb *chegar (arrive)* (o maravidil) post-positioned to the infinitive and, as it deals specifically with this verb, the preposition can, initially, be justified both by its regency (*chegar a algum lugar/ ponto (arrive at a place)*) as well as by the notion of restriction that it adds to the verb and that seems to be a determining factor to translate the notion of hierarchical order that the employment of the construction [CHEGAR + A + INFINITIVO] reflects. This specificity, however, does not explain the use of the preposition in the other constructions with unaccusative verbs, but rather shows that the underlying sense of the preposition concurs to determine the type of aspectual notion that the compound expresses, which corroborates with the argument that the preposition is not a meaningless item (Cf. LURAGHI, 2003).

However, it was in the analysis of the criteria adopted by the Spanish grammarian to affirm the stature of the prepositional constructions in the context of an auxiliary verb that we can reach an angle capable of shedding light upon the presence of the prepositions in the contexts of unaccusative verbs, such as those illustrated in (67) to (72), and to support the thesis we will explore below. By comparing the structure with the same meaning (*Empezó a llover* y *Empezó la lluvia; Terminó de llover y Terminó la lluvia)*, Bosque & Demonte (1999) argue that the noun phrase (*la lluvia*) cannot paradigmatically substitute the prepositional construction (*de llover*) as it acts as a subject of the verbs *começar* e *terminar (begin and end)*, a function that does not correspond to the infinitive, which, although it has a substantive function, is preceded by the preposition. Hence, it is possible to speculate that "the most notable innovation in the syntax of the infinitive"[53], according to Maurer Jr. (1959, p.185, our translation), have been a resource adopted in the constructions of the unaccusative auxiliary verb in the Romance languages to assert the formation of the compound, since the condition for auxiliarity is the presence of a single subject for the two verb forms, with the subject selected by the main form of the compound. In this sense, the infinitive governed by the preposition breaks, in the syntactic environment of the unaccusative verbs, as a mechanism that the linguistic system offers to the speaker to impede juxtaposed verbal forms from having two distinct subjects; consequently, they cannot be reanalyzed as an auxiliary verb construction. The choice of the preposition that will govern the infinitive,

---

[53] Original: "*a mais notável inovação na sintaxe do infinitivo*" (MAURER JUNIOR, 1959, p.185).

thus transforming it into the compound form capable of selecting the subject, will clearly be determined by its semantic potential to code the intended aspectual functions.

Therefore, this study demonstrates that the auxiliary verb is not responsible for the selection of the prepositional link that is interposed to the nominal form of the infinitive, which explains the possibility of the existence of constructions of indirect incidence regardless of the transitivity of the auxiliary verb form.

Final conclusions

The analysis reported herein chose as its object of study the verbal constructions of the Portuguese language composed by [auxiliary verb + preposition + infinitive]. The generalizations pointed out here result from the quantitative and qualitative analysis of the 802 diachronic pieces of data retrieved from the Tycho Brahe *Corpus* (http://www.tycho.iel.unicamp.br/corpus/), to create a linguistic sample from the fifteenth to the twentieth centuries. In addition, we qualitatively analyzed data from the *Portuguese Corpus* (https://www.corpusdoportugues.org/), together with those from our own intuition. Through a panchronic and multisystemic approach to language, we sought to answer four intimately related problem-questions, whose main conclusions can be summarized as follows:

(i)   Only four highly grammaticalized prepositions – A, DE, PARA, and POR – can occur in the context of the analyzed constructions, with DE being the most productive of these, answering for more than 90% of the occurrences from the *corpus*;

(ii)   Although some prepositions participated in the coding of more than one functionality, idiosyncrasies do exist: the prepositions A and POR participate only in the coding of aspect; the prepositions DE and PARA participate both in the coding of aspect and of modality; the preposition PARA is the only one that participates in the coding of time;

(iii) The elimination of the preposition in the context of auxiliary verb constructions of indirect incidence is limited to the preposition A and conditioned by a scaffolding of factors involving prosodic and phonological questions of the compound, in addition to the semantic content of the preposition;

(iv)   The subordinate tie that figures in the environment of auxiliary verb constructions results, in most of the constructions, from a context of purpose with the infinitive preceded by a preposition in which the complement is antepositioned. As this structure deals with constructions of unaccusative verbs, the presence of the prepositional link is a means through which to justify the formation of the compound, thus preventing the infinitive from being understood as the subject of the verbal form that precedes it.

Acknowledgements

COELHO, S. Um estudo de preposições em contexto de construções de verbo auxiliar. **Alfa**, São Paulo, v. 65, 2021.

■ *RESUMO: Neste artigo, descrevemos e analisamos construções da língua portuguesa formadas por [verbo auxiliar + preposição + infinitivo], que são aqui abordadas num viés pancrônico, pautado numa concepção cognitivista e multissistêmica de língua(gem). Nosso objetivo maior consiste em explicar por que, em algumas dessas construções, a preposição tende a ser apagada, sobretudo em contextos de oralidade, sem que tal apagamento resulte em prejuízo para a boa formação da sentença, enquanto em outras o liame preposicional deve obrigatoriamente se interpor entre o verbo auxiliar e a forma nominal de infinitivo, sob pena de se comprometer a boa formação do composto. No entorno dessa questão maior, buscamos ainda descrever que tipo de preposição ocorre nessas construções, a correlação entre o tipo de preposição e a função gramatical (tempo, aspecto, modalidade) codificada pela construção, além de identificar fatores condicionantes da presença da preposição nesses contextos. Os resultados obtidos acusam que apenas as preposições A, DE, PARA e POR podem ocorrer nesses contextos e que elas são herdadas do contexto de reanálise da construção, majoritariamente ligado a um ambiente sintático de finalidade. A opcionalidade da preposição no contexto é restrita à preposição A e condicionada por uma imbricação de fatores de natureza fonológica, prosódica e semântica.*

■ *PALAVRAS-CHAVE: Preposição. Construção de verbo auxiliar de incidência indireta. Aspecto verbal. Tempo. Modalidade.*

REFERENCES

ALMEIDA, João de. **Introdução ao estudo das perífrases verbais de infinitivo**. Assis: ILHPA-HUCITEC, 1980.

ALMEIDA, Napoleão M. de. **Gramática metódica da Língua Portuguesa**. 22.ed. São Paulo: Saraiva, 1969.

BARROSO, H. **O aspecto verbal perifrástico em português contemporâneo**: visão funcional/sincrônica. Porto: Porto Ed., 1994.

BASSETO, B. F. **Elementos de filologia românica**: história interna das línguas românicas. São Paulo: Edusp, 2010. v.2.

BERTUCCI, R. Aspecto terminativo: verbos auxiliares no português brasileiro. **Filologia e linguística portuguesa,** São Paulo, v.12, n.1, p.41-58, jan./jun. 2010.

BYBEE, J.; PERKINS, R.; PAGLIUCA, W. **The evolution of gramar**: tense, aspect and modality in the languages of the world. London: The University of Chicago Press, 1994.

BOSQUE, I.; DEMONTE, V. **Gramática descriptiva de la Lengua Española**: las construcciones sintácticas fundamentales: relaciones temporales, aspectuales y modales. Madrid: Espasa Calpe, 1999.

CAGLIARI, L. C. **Análise fonológica**: introdução à teoria e à prática com especial destaque para o modelo fonêmico. Campinas: Edição do Autor, 1997. pt. I.

CÂMARA JÚNIOR, J. M. **Dicionário de linguística e gramática**: referente à língua portuguesa. 28.ed. Petrópolis: Vozes, 2011.

CÂMARA JÚNIOR, J. M. **Problemas de linguística descritiva.** 17.ed. Petrópolis: Vozes, 1998.

CÂMARA JÚNIOR, J. M. **História e estrutura da Língua Portuguesa**. 2.ed. Rio de Janeiro: Padrão, 1976.

CASTILHO, A. T. de. **Nova Gramática do Português Brasileiro**. São Paulo: Contexto, 2010.

COELHO, S. M. Gradualismo do processo de gramaticalização e princípio da persistência: indícios de uma hierarquia de traços? **Filologia e Linguística Portuguesa,** São Paulo, v.15, n.2, p.519-541, jan./jun. 2013.

DAVIES, M.; FERREIRA, M. J. **Corpus do Português**: 45 milhões de palavras, 1300s a 1900s. 2006. [*S. l.*]: National Endowment for the Humanities, 2004-2015. Available in: http://www.corpusdoportugues.org/x.asp. Access on: 25 may 2021.

DIAS, A. E. da S. **Syntaxe histórica portuguesa**. Lisboa: Clássica, 1970.

GALVES, C.; ANDRADE, A. L. de; FARIA, P. **Tycho Brahe Parsed Corpus of Historical Portuguese**. [*S. l.*: *s. n.*, 2017). Available in: http://www.tycho.iel.unicamp.br/~tycho/corpus/texts/psd.zip. Access on: 25 may 2021.

GIVON, T. **Functionalism and gramar**. Amsterdam; Philadelphia: John Benjamins Publishing, 1995.

GIVON, T. **Syntax**: a functional- typological introduction. Amsterdam; Philadelphia: John Benjamins, 1984.

HEINE, B.; CLAUDI, U.; HÜNNEMEYER, F. **Grammaticalization**: a conceptual framework. Chicago: The University of Chicago, 1991.

HOPPER, P. On some principles of grammaticalization. *In*: TRAUBOTT, E. C.; HEINE, B.; (org.). **Approaches to grammaticalization**. Amsterdam: John Benjamins, 1971. p.17-36.

KURYLOWICZ, J. The evolution of grammatical categories. *In*: KURYLOWICZ, J. **Esquisses Linguistiques II**. München: Fink, 1965. p.38-54.

LEHMANN, C. **Thoughts on grammaticalization**: Munich: Lincom Europa, 1995. Originalmente publicado: Thoughts on grammaticalization: a programatic sketch. Köln: Arbeiten des Kölner Universalien 49 – Projects, 1982, v.1.

LURAGHI, S. **On the meaning of prepositions and cases**: the expression of the semantic roles in Ancient Greek. Amsterdam; Philadelphia: John Benjamins Publising, 2003.

MATTOS E SILVA, R. V. **O português arcaico**: fonologia. 4.ed. São Paulo: Contexto, 2001.

MAURER JÚNIOR, T. H. **Gramática do latim vulgar**. Rio de Janeiro: Livraria Acadêmica, 1959.

PONTES, E. **Espaço e tempo na Língua Portuguesa.** Campinas: Pontes, 1992.

SAID ALI, M. **Dificuldades da língua portuguesa**. 7.ed. Rio de Janeiro: ABL: Biblioteca Nacional, 2008. (Coleção Antônio de Morais Silva, v.7).

SAID ALI, M. **Gramática histórica da Língua Portuguesa**. São Paulo: Melhoramentos, 1966.

SANTANA, L. **Relações de complementação no português brasileiro**: uma perspectiva discursivo-funcional. São Paulo: Cultura Acadêmica, 2010.

SOUSA, N. dos S. V. **Preposições em perífrases verbais de infinitivo**: um estudo preliminar. 2011. 93f. Trabalho apresentado à disciplina Projeto de Curso. Instituto de Letras, Universidade de Brasília, Brasília, 2011.

TEYSSIER, P. **História da língua portuguesa**. Trad. Celso Cunha. São Paulo: Martins Fontes, 1997.

TRAVAGLIA, L. C. A (poli)gramaticalização do verbo *acabar*. **Letras & Letras**: Uberlândia, v.20, n.2, p.21-56, jul./dez. 2004.

TRAVAGLIA, L. C. **O aspecto verbal no português**: a categoria e sua expressão. Ed. rev. Uberlândia: Universidade Federal de Uberlândia, 1985.

VENDLER, Z. **Verbs and times:** Linguistics in Philosophy. Ithaca: Cornell University Press, 1967.