

Repositório ISCTE-IUL

Deposited in *Repositório ISCTE-IUL*:

2025-07-03

Deposited version:

Publisher Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Bico, M. I., Cardeira, E., Baptista, J. & Batista, F. (2024). Support verb constructions in medieval Portuguese: Evidence from the CTA corpus. In Pablo Gamallo, Daniela Claro, António Teixeira, Livy Real, Marcos Garcia, Hugo Gonçalo Oliveira, Raquel Amaro (Ed.), Proceedings of the 16th International Conference on Computational Processing of Portuguese: Vol.2. (pp. 118-129).: Association for Computational Linguistics.

Further information on publisher's website:

<https://aclanthology.org/2024.propor-2.16>

Publisher's copyright statement:

This is the peer reviewed version of the following article: Bico, M. I., Cardeira, E., Baptista, J. & Batista, F. (2024). Support verb constructions in medieval Portuguese: Evidence from the CTA corpus. In Pablo Gamallo, Daniela Claro, António Teixeira, Livy Real, Marcos Garcia, Hugo Gonçalo Oliveira, Raquel Amaro (Ed.), Proceedings of the 16th International Conference on Computational Processing of Portuguese: Vol.2. (pp. 118-129).: Association for Computational Linguistics.. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Support Verb Constructions in Medieval Portuguese: Evidence from the CTA Corpus

M.^a Inês Bico^{ib}, Esperança Carneira^{ib} Jorge Baptista^{ib}

Univ. Lisboa
Faculdade de Letras
Centro de Linguística
Lisboa, Portugal
mariainesb1@edu.ulisboa.pt
ecarneira@edu.ulisboa.pt

Univ. do Algarve
Faro, Portugal
INESC-ID Lisboa,
Portugal
jbaptis@ualg.pt

Fernando Baptista^{ib}
ISCTE, Inst. Univ. Lisboa
INESC-ID Lisboa, Portugal
fernando.batista@iscte-iul.pt

Abstract

This paper conducts a systematic survey of Support Verb Constructions (SVC) in Medieval Portuguese using the Corpus de Textos Antigos (CTA) corpus. SVC involves verb-noun combinations, where the noun serves as the main predicative element, and the verb conveys grammatical values. Limited historical evidence exists for SVC in earlier stages of Portuguese, with previous studies lacking digital access and Natural Language Processing tools. This study utilizes a subset of the CTA, comprising around 0.5 million tokens, annotated for part-of-speech and lemmata. Out of 175 candidate support verbs in Modern Portuguese, 81 were identified in the corpus, totaling 27,645 occurrences. Manual analysis of concordances revealed a little more than 3,000 SVC instances and more than 900 predicative nouns (types), uncovering several linguistic phenomena. The paper delineates the adopted procedure and explores essential linguistic properties of identified SVC in the CTA corpus, emphasizing the significance of leveraging NLP tools for a comprehensive linguistic description of Old Portuguese.

1 Introduction

This paper presents a systematic survey of *support verb constructions* (SVC) from Old Portuguese (OP) and Middle Portuguese (MP) found in a corpus. Old Portuguese is an early stage of the Portuguese language that was written from the 12th century to the late 14th century or early 15th century, when Middle Portuguese starts. This period of the language goes until the mid-16th century (Castro, 2006).

Support verb constructions (Harris, 1964, 1991; Gross, 1981, 1996) are elementary sentences, consisting of verb-noun combinations where the main predicative element is the noun, called a *predicative noun* (*Npred*), and the verb, called a *support verb* (*Vsup*), has mostly a grammatical auxiliary

function, conveying the tense and person-number agreement values.

Next, there is a clear example of a SVC.¹

(1) [OP] *e nõ êtendedes quanta ãJuria fazedes ao vosso criador* [mPT] *‘e não entendeis quanta injúria fazeis ao vosso criador’* [EN] *‘and you do not understand how much injury (= insults, offenses) you do to your creator.’* [HdE-A:142892]

In this instance extracted from the ms. A of *Horto do Esposo* (HdE-A), the noun *ãJuria* ‘*injúria*’ ‘*injury*’ is supported by the verb *fazedes* ‘*fazeis*’ ‘*do-present-2nd-person-pl.*’. The syntactic structure, as determined by the predicative noun, reveals a *dative* complement *ao vosso criador* ‘*idem*’ ‘*to your creator*’, along with the distributional constraints placed on the human subject. The dative type of complement can be further substantiated by examining other occurrences of the noun in the corpus.

(2) [OP] *lhe perdõ das ãJurias que lhe fezera* [mPT] *‘lhe perdoou das injúrias que lhe fizera’* [EN] *‘forgave him/her for the insults that had been done to him/her’* [HdE-A:140338]

The identification of the SVC involving this noun in Old Portuguese texts is facilitated not only by its persistence in Modern Portuguese (albeit with potentially nuanced differences in meaning) but also by its consistent display of the same syntactic construction and distributional constraints. However, in this context, particular em-

¹Conventions: In the examples, a color code indicates the language: *purple* for *Old Portuguese* (OP), *red* for the *Modern Portuguese* (mPT) gloss, and *blue* for an *English* gloss (EN); occasionally, *gray* will be used for *Modern Portuguese* (PT) example. For clarity, examples and glosses are also preceded by the code of each language. Targeted words, usually the support verb and the predicative noun are highlighted in bold typeface. The alphanumeric codes at the end of the examples identify the text and the token number where the verb occurred in the corpus. The list of texts in the corpus considered in this paper is presented in (3.1). In accordance with the transcription standards adopted for the corpus, words are presented in their original spelling, including the use of uppercase/lowercase letters as well as diacritical marks.

phasis should be placed on the observations put forth by Ranchhod (1999, p. 3) regarding the use of corpora in studying the earlier stages of language evolution:

When studies are conducted on earlier stages of the evolution of a language, the knowledge derived from it comes from what is evident in the texts of the period or periods under investigation. Nothing that is not present in the texts can be presumed; it is not possible to make judgments of acceptability about constructions and usages that are not attested there. Only the analysis of these data and theoretically grounded argumentation can fill the gap in these refutation criteria. (our translation)

2 Related Work

The study of support-verb constructions in old stages of Portuguese, namely Medieval Portuguese, is scarce. The two principal works are those of Chacoto (1996) and Ranchhod (1999), both studying the existence of these constructions in different texts. Though different corpora were used, the conclusions are similar: i) SVC are a patrimony of the Portuguese language, going as far back as the medieval period; ii) there are a lot of similarities between Medieval and Modern Portuguese. The more significant difference noted by both authors are the position of the constituents of SVC: in certain types of constructions, not possible today, Medieval Portuguese permitted the placement of the predicative noun before the support verb. For example, Ranchhod (1999) identifies this different word order in *Crónica Geral de Espanha de 1344* [‘General Chronicle of Spain from 1344’].

(3) [OP] *despois que ambas amor ouvemos*
[mPT] *‘depois que ambas amor houvemos’*
[EN] *‘after we both love had’*

Chacoto (1996) studies SVC with the verb *fazer* ‘to do’, in *Demanda do Santo Graal* [‘The Quest of the Holy Grail’]. The author distinguishes three uses of the verb: full or distributional verbs, causative operator-verbs e support-verb constructions. Following a brief description and characterization of these types of constructions, the focus turns to cases wherein exists nominalization (‘*fazer pecado*’= ‘*pecar*’ ‘to do a sin’ = ‘to sin’), giving also examples for when this relation does not exist and what reasons may explain this.

Ranchhod (1999) in its turns studies SVC in fifteen chapters of *Crónica Geral de Espanha de 1344*, noting the high frequency of these constructions. Defining the SVC in an elementary phrase, the author also discusses SVC as nominal groups, nominalizations and autonomous predicative nouns. The most frequent verbs were found to be ‘*haver*’ ‘have’ and ‘*fazer*’ ‘do’, while the verbs ‘*ser*’ and ‘*estar*’ ‘be’ are not yet used in SVC.

More recently, Pante and Ortega (2009) and Pante (2012) studied SVC with the verb ‘*tomar*’ ‘take’ in different medieval texts. Pante and Ortega (2009) uses two late 15th century texts (*Leal Conselheiro* [‘Loyal Adviser’] and *Livro da Enseñança de bem cavalgar toda a sela* [‘Book of the lesson on how to ride well every saddle’]) to shine light on the reasons why SVC are used: to allow the characterization of the predicative noun and to reduce the valence of the verbs. Pante (2012) studies this verb in *Horto do Esposo*, which is also present in our corpus. However, only instances in which it is possible to discern a case of nominalization were included, in order to identify the possible reasons for choosing a SVC instead of a distributional verb. Nine pairs of SVC and nominalizations were found, and results were similar to those already mentioned above.

3 Methods

This paper utilizes the part-of-speech and lemmatized partition of the CTA corpus to conduct a systematic survey of Medieval Portuguese SVC. In this Section, we elaborate on the methodology employed.

3.1 The CTA Corpus

The Old Portuguese Texts Corpus (Corpus de Textos Antigos - CTA)² comprises 32 texts (as of January 2024) written until 1525. Adhering to a principle of high fidelity to the source documents, the editions maintain the original graphic forms, lacunae, and errors. Utilizing the web-based framework TEITOK (Janssen, 2016; Vaamonde and Janssen, 2020), the corpus combines text with linguistic annotation.

Established in 2015, the CTA corpus has been subject to a data enhancement project, employing Natural Language Processing tools and meth-

²<https://teitok.clul.ul.pt/teitok/cta/index.php?action=home>

ods (Bico et al., 2022). One of the first tasks done was the enrichment of the data with lemmata and part-of-speech tags, with the respective morphosyntactic category and inflexion values. As all original graphic forms are maintained in the corpus, graphic variance poses a challenge for search queries. This is overcome by the existence of lemmas aggregating different forms and granular POS tags which allow the optimization of sophisticated lexical queries. For SVC, filtering both lemmatized nouns and verbs enables a more extensive cross-search spanning the entire corpus. For instance, in examples (4)-(6), both the noun *avito* and *religiã* exhibit various graphic forms.

(4) [OP] *tomasse ho avyto da Relligjã* [mPT] ‘*tomasse o hábito da religião*’ [EN] ‘[he/she] would take the habit of religion’ [MISJ:17755]

(5) [OP] *tomou ho nosso sancto avito da Religiã* [mPT] ‘*tomou o nosso santo hábito da religião*’ [EN] ‘[he/she] took our holy habit of religion’ [MISJ:21060]

(6) [OP] *tomara ho avito da Relligiam* [mPT] ‘*tomara o hábito da religião*’ [EN] ‘[he/she] had taken the habit of religion’ [MISJ:23066]

As described in (Bico et al., 2022), after a manual annotation task of the 150K tokens that make the ms. A of *Horto do Esposo* (HdE-A), two automatic annotation experiments were done using TreeTagger (Schmid, 1994, 1999). In the first experiment, a model trained with the batch of 150K tokens was applied on a set of two new texts: fragments D, C and E of *Horto do Esposo* (HdE-DCE), and ms. G1 of *Vida e Milagres de Santa Senhorinha de Bastos* (VMSSB-G1). The results varied between 91% and 67% of precision. Incorrectly annotated data was then manually revised and fed to a new, second model, which now comprised c. 167K tokens for training. The second experiment, which followed the premises of the first, applied the second model to four new, larger texts: *História de mui nobre Vespasiano*, *Memorial da Infanta Santa Joana*, *Livro dos Mártires*, and ms. E of *Vida e Milagres de Santa Senhorinha de Bastos* (respectively, Vespasiano, MISJ, LdM, VMSSB-E). Results slightly increased, varying between 77% and 82%. The manual revision task is of the utmost importance as it ensures the data correction and refinement for following annotation models and further work with the data. In the end, the corpus currently has seven annotated texts (one manually, six semi-automatically), whose combined totals

Text	Tokens	Precision
HdE-A	154,891	
HdE-DCE	2,694	90.97%
VMSSB-G1	13,944	67.41%
VMSSB-E	14,747	76.96%
Vesp[asiano]	19,134	82.36%
LdM	253,277	80.45%
MISJ	51,679	77.22%
Total	510,366	

Table 1: Total of annotated tokens in CTA and precision rates of the automatic annotation.

surmounts more than half a million of annotated tokens (Table 1).

3.2 Selected Texts

Horto do Esposo was first written in Portuguese, at the turn of the 14th to the 15th centuries (1390-1437), in the *scriptorium* of Mosteiro de Santa Maria de Alcobaça. Manuscript HdE-A dates from the same period. HdE-DCE is made of three fragments from a manuscript from Mosteiro do Lorvão. Spiritually didactic, the text uses *exempla* to illustrate several subjects, often recurring to the *auctoritas* and the Holy Bible.

Of the *História de mui nobre Vespasiano* there is only one known witness: the 1496 incunable printed in Lisbon by Valentim Fernandes. It tells the novelistic story of Roman emperor Vespasian who, to be cured from leprosy, seeks relics from Jerusalem.

The *Livro dos Mártires* was first written in Castilian by Bernardo de Briuega, at the mandate of King Alfonso X of Castille. It is the third part of a five book project and deals with the life of saints and martyrs from the beginning of Christendom. The text was translated into Portuguese during the reign of King Denis of Portugal (1279-1325), and it is known that a 15th century manuscript existed. However, the only surviving complete witness is the 1513 incunable printed in Lisbon by João Pedro Bonhominy.

The *Memorial of Infanta Santa Joana* tells the story of Infanta Joana of Portugal (1452-1490), daughter of King Afonso V of Portugal, and her entrance and life in the Mosteiro de Jesus of Aveiro. The text is supposed to have been first written between 1513 and 1525, and the copy is posterior to 1525.

The *Vida e Milagres de Santa Senhorinha de Basto* is a hagiographic tale of the life of Saint Senhorinha, a Portuguese abbess from the 10th century. There are four known witnesses of this text, all present in CTA, but only VMSSB-G1 and VMSSB-E have been lemmatized and annotated so far. The text was first written in the 13th century (1248-1284); ms. G1 was produced in the 17th century (1620-1645) and ms. E between 1692 and 1705.

3.3 Corpus processing

From an existing (though non-exhaustive) list of 175 support verbs previously identified in Modern Portuguese (Baptista and Mamede, 2020a,b), concordances were automatically extracted from the annotated partition of the corpus. The verbs on the list are not exclusively used for SVC and can also be used as full verbs, auxiliary verbs, etc. A 7-word window was applied for both the left and right context. A list of 27,645 instances of these verbs was automatically retrieved for 81 verbs of that initial list. Then, a manual identification of SVC was made, distinguishing them from other uses of these verbs. In this analysis, the predicative noun is indicated next to the respective concordance, along with the signs < and >, depending on whether the noun appears to the left or to the right of the support verb. Table 2 displays the distribution of some of these candidate support verbs, simultaneously, the most and the least frequently occurring, along with the number of SVC identified. A total of 2,982 SVC have been identified. This number is, in fact, slightly higher, as predicative nouns often appear coordinated, e.g. *fe e Confyāca* ‘fé e confiança’ ‘faith and confidence’. A total of 112 instances of this coordinated noun phrases have been found, raising the total to 3,094 SVC found (949 different types). For 2,423 instances, the predicative noun is on the right of the support verb, and for 559 are on the left (often as antecedent of a relative subclause). To the best of our knowledge, this is the largest collection of SVC in Medieval Portuguese available.³ The identification of SVC is still ongoing: verbs ‘ser’ (11,153 occurrences) and ‘estar’ (1,424 occurrences) ‘be’ were not considered at this stage.

The most frequently occurring Vsup verbs are ‘haver’ and ‘fazer’ (‘to have’, ‘to do’), with

Verb	Total	SVC
haver (<i>have</i>)	4,163	1,556
fazer (<i>do</i>)	3,613	1,216
ter (<i>have</i>)	1,117	61
tomar (<i>take</i>)	667	91
...		
usar (<i>use</i>)	105	20
reinar (<i>reign</i>)	85	4
falecer (<i>die</i>)	83	4
cobrir (<i>cover</i>)	74	7
mover (<i>move</i>)	73	1
manter (<i>maintain</i>)	32	7
cessar (<i>cease</i>)	31	3
cometer (<i>commit</i>)	22	3
firmar (<i>firm</i>)	8	1
conceder (<i>concede</i>)	5	2
incorrer (<i>incur</i>)	5	5
render (<i>yield</i>)	5	3
gozar (<i>enjoy</i>)	3	2
proferir (<i>utter</i>)	1	1

Table 2: Distribution of SVC in the corpus

1,470 and 1,243 instances as Vsup in SVC, respectively. While not reaching these high frequencies, it’s worth noting that the verbs ‘ter’ and ‘tomar’ (‘to have’, ‘to take’) also appear, with 61 and 91 occurrences, respectively. All other verbs exhibit a residual presence in the corpus, contributing only a few instances of SVC each.

4 Results: SVC in the CTA

Not being possible to analyse all of the cases found, only note-worthy observations will be made in this Section.

4.1 Nominalizations

Numerous SVC, featuring predicative nouns, find their equivalent in transformationally related verbal constructions, as noted by various scholars (Harris, 1964, 1991; Gross, 1981). Moreover, these instances of *nominalization*, a paraphrastic equivalence between the full verb and the predicative noun constructions, exhibit identical meanings and distributional constraints. Consequently, they hold a transformational status within the grammatical structure (Harris, 1964, 1991).

This phenomenon is also apparent in Old Portuguese, as the corpus presents instances of SVC alongside the corresponding verbal predicates. An illustrative example is *alguém fazer injúria a al-*

³Table 3 in the Appendix A shows a sample of the concordances with SVC. Appendix B shows the frequency of the predicative nouns types.

guém = *alguém injuriar alguém*, which translates to ‘someone doing/giving offense/insult to someone’ = ‘someone offending/insulting someone’. The SVC has already been shown in examples (1)–(2). Additionally, there is evidence of the corresponding verbal construction in (7), which even encompasses a passive construction in (8).

(7) [OP] *alugue quẽ me ẽJuriasse*. [mPT] ‘alguém que me **injuriasse**.’ [EN] ‘someone who would insult/offend me’ [HdE-A:95923]

(8) [OP] *enfadou de seer ẽJuriado per cinco ãos* [mPT] ‘**enfadou-se de ser injuriado por cinco anos**’ [EN] ‘became fed up with being insulted/offended for five years’ [HdE-A:95976]

4.2 Standalone predicative nouns

However, the presence of a nominalization is not a prerequisite for establishing the predicative nature of a noun. Indeed, numerous predicative nouns exist as morphologically isolated lexical units, lacking any corresponding verbal or adjectival equivalent. These *standalone* or *autonomous predicative nouns* (Gross, 1981; Ranchhod, 1990; Baptista, 2005; Baptista and Mamede, 2020b), can be identified through various means: (i) their retention of the same function in Modern Portuguese, often accompanied by a similar syntactic construction; this is the case of the noun *ẽJuria* ‘**injúria**’ ‘insult/offense’, in the examples above; (ii) there is supporting evidence from corpus, showcasing syntactic patterns alongside other predicative nouns from the same text; this can be seen in example (9):

(9) [OP] *ella tijnha fe e Confyãca ẽ deus* [mPT] ‘**ela tinha fé e confiança em deus**’ [EN] ‘she had faith and trust in God’ [MISJ:26966]

In this example, the predicative nouns *fe* ‘fé’ ‘faith’ and *Confyãca* ‘**confiança**’ ‘trust’ appear coordinated under the support verb *tẽer* ‘**ter**’ ‘have’, with the same subject *ella* ‘**ela**’ ‘she’ and the same complement *deus* ‘**deus**’ ‘God’. This example allows one to consider both nouns as predicates expressed by SVC. On one hand, both nouns persist in Modern Portuguese, maintaining the same construction and conveying similar meanings, as per criterion (i). On the other hand, the coordination of these nouns under the same support verb in (5) implies, in accordance with criterion (ii), that they also serve a similar function. Indeed, *Confyãca* serves as a predicative noun, evident in its standalone occurrences with the same support verb, as in (9). Additionally, there is evidence of the cor-

responding nominalization in texts from the same period, as seen in (10)–(11).

(10) [OP] *pois confiaça tẽ em deos* [mPT] ‘**pois confiança tem em deus**’ [EN] ‘for [s/he] has trust in God’ [LdM:27891]

(11) [OP] *e todosos que cõfiam ẽ elle* [mPT] ‘**e todos os que confiam nele [= em_ele]**’ [EN] ‘and all those who trust in him.’ [HdE-A:48062]

In turn, in spite of not having a verbal equivalent construction, the occurrence of *fe* (or the variant *fee*) ‘fé’ ‘faith’ both in sentences like (9), coordinated with another predicative noun, and in standalone mode in sentences like (12), allows one to consider *fe* a predicative noun and *tẽer* ‘**ter**’ or *aver* ‘**haver**’ ‘have’ its support verbs.

(12) [OP] *e a fe que eles tijnhã* [mPT] ‘**e a fé que eles tinham**’ [EN] ‘and the faith that they had.’ [LdM:216921]

In certain cases, the support verb construction (SVC) has disappeared from use, despite the continued presence of the individual words—both the support verb and the noun—in the modern language. This is the case of the SVC *ter mentes* (13):

(13) [OP] *Hũũ sabedor teue mẽtes ẽnas cousas do mũdo* [mPT] ‘**Um homem sábio teve mentes nas coisas do mundo**’ [EN] ‘A wise man had minds on the things of the world/on worldly matters’ [HdE-A:57369]

whose meaning could be defined as ‘to consider, to reflect upon’ or similar expressions. A systematic analysis of the usage of this expression, facilitated by the annotated corpus used here, can contribute to a more precise understanding of both its meaning and syntactic properties.

4.3 Variant support verbs

One of the defining properties of SVC is that the same predicative noun can select, besides its elementary *Vsup*, the one with the least semantic load and broader distribution, a variety of other support verbs, which introduce nuances of aspect (Fotopoulou et al., 2021), modality and style. This variation is also abundantly exemplified in the corpus, as seen in examples (14a)–(14d) with the noun *fe* ‘fé’ ‘faith’. This nouns displays several SVC both with the elementary support verbs *tẽer* ‘**ter**’ ‘have’ (14a) and ‘**aver**’ ‘**haver/ter**’ ‘have’ (14b); the durative/permansive variant *manter* ‘**id.**’ ‘keep/maintain’ (14c); and the terminative variant *perder* ‘**id.**’ ‘lose’ (14d), respectively.

(14a) [OP] *a ajuda de deus . ã que tinha toda sperãca e ffe* [mPT] ‘a ajuda de Deus, em quem tinha toda a esperança e fé’ [EN] ‘help of God, in whom he had all hope and faith’ [MISJ:26289]

(14b) [OP] *por sempre fostes mentirosos e que no auedes fee nemhuia* [mPT] ‘porque sempre [vós] fostes mentirosos e porque não haveis fé nenhuma’ [EN] ‘because [you_pl.] have always been liars and because [you_pl.] have no faith.’ [LdM:93832]

(14c) [OP] *E o que mantem aquesta fe pola boã obra sera parceiro* [mPT] ‘E o que mantém esta fé pela boa obra será parceiro’ [EN] ‘And the one who maintains this faith through good deeds will be a partner’ [LdM:64555]

(14d) [OP] *homẽs maos e desleaes que auia perdida a sua fe.* [mPT] ‘homens maus e desleais que tinham perdido a sua fé’ [EN] ‘evil and disloyal men who had lost their faith’ [LdM:193826]

These examples do not exhaust the full syntactic diversity of the SVC involving the noun *fé* and other constructions in which this noun appears in the corpus.⁴

As mentioned in the literature, we also observed a differential distribution between the more commonly used *aver* and the less frequent *teer*. This

⁴Due to space constraints, we cannot explore them further, but we can briefly mention a few, providing a tentative description: (i) Permansive SVC ‘*ter-se na fé*’:

(15) [OP] *huã martire que se teve firmemente na fé de nosso senhor* [mPT] ‘um mártir que se manteve firmemente na fé de nosso senhor’ [EN] ‘a martyr who remained steadfast in the faith of our Lord’ [LdM:296]; (ii) Terminative SVC with the predicative noun as the subject:

(16) [OP] *Mas eu roguei por ti que tua fé não falhasse* [mPT] ‘Mas eu roguei por ti [para] que a tua fé não falhasse’ [EN] ‘But I have prayed for you [so] that your faith may not fail’ [LdM:23046]; (iii) Another SVC with a kind of echo complement introduced by *em*:

(17) [OP] *Se ouverdes em vós fé tanto como grão de sebe* [mPT] ‘Se tivéreis em vós tanta fé como grão de sebe’ [EN] ‘If you had within you so much faith like a grain of mustard seed’ [LdM:74610]; (iv) An (arguably) SVC with *ficar em*:

(18) [OP] *E muitos ficaram em na fé de nosso senhor Jesus Cristo* [mPT] ‘E muitos ficaram na fé de nosso senhor Jesus Cristo’ [EN] ‘And many remained in the faith of our Lord Jesus Christ’ [LdM:65467]; (v) Another (arguably) SVC with *tornar-se a*:

(19) [OP] *E aconselhavam-lhes estes santos todo o dia que se tornassem à fé de nosso senhor Jesus Cristo* [mPT] ‘E aconselhavam-lhes estes santos todo o dia que regressassem à fé de nosso senhor Jesus Cristo’ [EN] ‘“And these saints advised them every day to return to the faith of our Lord Jesus Christ.”’ [LdM:68776]; (vi) Finally, an autonomous (unrelated) SVC with *dar*, still existing in Modern Portuguese: (20) [OP] *E eu dando fee aas palauras delles* [mPT] ‘E eu dando fé às palavras deles’ [EN] ‘And I, giving credence to their words’ [LdM:29731]. Other constructions exist but cannot be further detailed.

trend may have eventually resulted in ‘*ter*’ assuming the role of support verb that the former held.

4.4 Conversion

A distinctive type of SVC involves *Conversion*, a transformation introduced by Gross (1981) and extensively described for Modern French by Gross (1989, 2022). This transformation is akin to a verbal active/passive transformation, but is specific to SVC. During this operation, the positions of the arguments of the predicative noun are altered: the agent-subject in the standard, active-like SVC becomes the prepositional complement in the converse, passive-like SVC. Simultaneously, the patient/object-complement becomes the grammatical subject of the converse, and the support verb in the standard construction is replaced by a converse support verb.

This transformation has been extensively examined in Modern Portuguese Ranchhod (1990); Baptista (2005); Chacoto (2005); Rassi et al. (2016); Rassi (2023), among others. Evidence of it is also present in the CTA corpus, as illustrated by examples (21)–(22), with *mercee* ‘*mercê*’ ‘*favor/gift*’:

(21) [OP] *Fazé me tanta mercee* [mPT] ‘Fazei-me tanta *mercê*’ [EN] ‘Do me such a favor’ [MISJ:26593]

(22) [OP] *eu requebendo delle mercee* [mPT] ‘eu recebendo dele *mercê*’ [EN] ‘I receiving from him [a] favor’ [LdM:230811]

However, it’s important to note that the same predicative noun can determine multiple SVCs, each characterized by different support verbs and structures, and denoting distinct meanings. An example of this is the noun *mercee*, which conveys a different notion of ‘ *piedade*’ ‘*pity/mercy*’ in another SVC, as demonstrated in (23) with the support verb *aver* ‘*haver/ter*’ ‘*have*’.

(23) [OP] *bõa mulher aue merçe de mi* [mPT] ‘boa mulher, tende *piedade de mim*’ [EN] ‘good woman, have mercy on me’ [LdM:80162]

The attentive reader would likely have observed that, in these examples, the noun appears in three different spellings. Consequently, these nuanced distinctions can only be discerned through a systematic examination of the distribution of a given noun in the corpus. This is made feasible by querying it through the same (modern) lemma, ‘*mercê*’, as it was done here.

4.5 Operator-verb constructions

During the process of identifying SVCs, care is taken to avoid confusing support verbs with other syntactic constructions. Indeed, certain verbs, besides their role as support verbs, display a range of syntactic constructions, including both *full* or *distributional verbs* (24) and (tense) *auxiliary verbs* (25).

(24) [OP] *Como sam sadornim mādou fazer hũa casa muy pequena açerca da egreja* [mPT] ‘*Como São Sandornim mandou fazer uma casa muito pequena próximo da igreja*’ [EN] ‘*How Saint Sandornim ordered to have a very small house built near the church.*’ [LdM:10525]

(25) [OP] *e que lhe tinham os inimigos cercado o castello d aguiar*⁵ [mPT] ‘*e que os inimigos lhe tinham cercado o castelo de Aguiar*’ [EN] ‘*and that his enemies had surrounded the castle of Aguiar*’ [VM-G1:11353]

In this context, and due to its potential confusion with SVC, this paper emphasizes a specific verb-noun construction: *operator-verb constructions* (Vop). Operator-verbs, a theoretical construct introduced by Gross (1981), serve to describe a complex construction formally resembling SVC but incorporating a new element of **cause**, while maintaining the distributional constraints between the predicative noun and its arguments. Consequently, they are also referred to as *causative operator-verbs* (Vopc). To be precise, Vopc operate on an elementary sentence, introducing a new element and establishing a causal relationship between this element and the sentence they modify. Typically, Vopc induce some form of restructuring in the sentence, normally ‘absorbing’ the support verb and altering the syntactical function or case of the predicate noun and its arguments within the sentence. These operations may result in surface sequences where the Vopc and the predicative noun formally resemble those of SVC. Example (26) provides an illustration of the Vopc *fazer* ‘id.’ ‘do/make’ operating on the predicative nouns *temor* ‘id.’ ‘fear’ and *espanto* ‘id.’ ‘astonishment’.

(26) [OP] *faziam me temor e graue espanto* [mPT] ‘*faziam-me temor e grave/muito espanto*’ [EN] ‘*[they] filled me with fear and great astonishment*’ [HdE-A:32599]

⁵We do not enter here in the discussion of this periphrastic verbal tense and the agreement of the part participle with the sentence subject. Please see below the concept of Vopl.

The SVC involving *temor* and *espanto* are otherwise well-documented in the corpus (12)-(13), employing the verb *haver* ‘have’, which would be replaced in Modern Portuguese by *ter* ‘have’, as seen from examples (27)-(28):

(27) [OP] *emristeçer e hauer temor e angustia* [mPT] ‘*entristecer e ter temor e angústia*’ [EN] ‘*to sadden/become sad and to have fear and anguish*’ [LdM:23327]

(28) [OP] *grande foy o espāto que ouue porque* [mPT] ‘*grande foi o espanto que teve porque*’ [EN] ‘*great was the astonishment that he/she had because*’ [LdM:248295]

Another complex construction introduced by Gross (1981) is the *linking operator-verb* (Vopl). In this scenario, an element also is added to an elementary sentence, but this element is already an argument of the predicative noun, hence the term *linking*. This is indicated by a constraint of coreference with the predicative noun, typically accompanied by a possessive determiner that is obligatorily coreferent to the Vopl subject. Vopl constructions can thus be seen as a linguistic device for restructuring an SVC elementary sentence. This is achieved by extracting the linked element and placing it in the more prominent syntactic position of the subject in the sentence. Example (29) illustrates a Vopl construction.

(29) [OP] *e aquele entenda que tem a sua nobreza tēyera*. [mPT] ‘*e aquele entenda que tem a sua nobreza inteira*’ [EN] ‘*and let that person understand that he has his nobility intact*’ [HdE-A:78625]

In this example, *nobreza* is modified by a secondary predication, *tēyera* ‘intact’, and there is constraint coreference between the subject of *tēer* ‘ter’ ‘have’ and the possessive determiner of the noun (*sua* ‘id.’ ‘his’). In Modern Portuguese a similar Vopl construction exists, as shown by the gloss. However, in Medieval Portuguese, the noun *nobreza* selects the support verb *haver* ‘id.’ and not *tēer* (30). Similarly, in this stage of the language, the copula verb selected by the adjective *tēyero* is *seer* ‘ser’ ‘be’ and not *estar* ‘id.’ ‘be’ (31).

(30) [OP] *que ha grande nobreza do teu linhagem* [mPT] ‘*que tem a grande nobreza da tua linhagem*’ [EN] ‘*that has the great nobility of your lineage*’ [LdM:242119]

(31) [OP] *ca nō es tēyero ãno spiritu* [mPT] ‘*pois [tu] não estás inteiro no espírito*’ [EN] ‘*for [you] are not whole in spirit*’ [LdM:144083]

In this case, one could arguably analyze (31) as a *Vopl* construction of *ter* on the adjectival construction of *ẽteyro*, which predicates on the noun *nobreza*, as illustrated by the following structure (32):⁶

(32) [OP] *aquele_i tem # a nobreza de_aquele_i ẽteyra* [mPT] ‘*aquele_i tem # a nobreza de_aquele_i ẽ inteira*’ [EN] ‘that person has # the nobility of that person is intact’

In spite of having identified many of these instances of *Vopc* and *Vopl*, they were not considered at this stage of the annotation, and were left for future work.

4.6 Contrasting constructions

Several impersonal constructions, expressing *meteorological* predicates were identified (33)–(34).

(33) [OP] *Porque fazia entom frio* [mPT] ‘*Porque fazia entãõ frio*’ [EN] ‘because it was cold’ [LdM:24447]

(34) [OP] *e tu poseste as tebras e fez sse a noyte* [mPT] ‘*e tu puseste as trevas e fez-se a noite*’ [EN] ‘and you put the darkness and night was made’ [LdM:48130]

These constructions closely resemble those found in Modern Portuguese. The grammatical status of verbs such as ‘*fazer*’ ‘*do*’ is unclear, as well as its two contrasting constructions: the pronominal structure of this verb with *noyte* ‘*noite*’ ‘night’, against the non-pronominal structure with *frio* ‘*id.*’ ‘cold’.

Additional contrasting constructions were identified. Certain verbs, such as *ministrar* ‘*ministrate*’, do not seem to be utilized as support verbs in Medieval Portuguese. Out of the 13 occurrences of this verb, one can find as its direct complement the generic noun *cousas* ‘*coisas*’ ‘things’ (5 instances) and with the noun *riquezas* (1 instance), in the sense of ‘administrate’ (35):

(35) [OP] *que auia satẽta e tres que lhe menis-trauã suas riquezas . e suas herdades* [mPT] ‘*que havia sete e trẽs [anos] que lhe ministrava suas riquezas e suas herdades*’ [EN] ‘that there were seven and three [years] that he/she administered his/her wealth and his/her estates.’ [LdM:116992] The so-called ‘relative subclause without explicit antecedent’ (Veloso, 2013), has also been found (1 instance), for example (36):

(36) [OP] *cõ a dita Senhora steuesse e lhe minystrasse e aparelhasse o que avija de comer* [mPT] ‘*com a dita Senhora estivesse e lhe ministrasse e aparelhasse o que houver de comer*’ [EN] ‘that he/she be with the said Lady and serve and prepare what there is to eat’ [MISJ:18795].

This full/distributional verb ‘*ministrar*’ is no longer in use in Modern Portuguese. Instead, the verb *administrar* presents a very similar construction. This later verb also often occurs as a variant *Vsup* of *ministrar*. It is also worth mentioning an apparently intransitive construction, eventually with a dative beneficiary complement (37):

(37) [OP] *E logo sse ella leuãtou sãã e mynjstrou a todos.* [mPT] ‘*E logo se ela levantou curada, serviu a todos*’ [EN] ‘And she immediately got up, [fully] healed, and she served [to] everyone’ [LdM:78346]

In contrast, in Modern Portuguese SVC, *ministrar* is a relatively common stylistic variant of *dar*, as seen in examples like *ministrar* – *assistência*, *curso*, *medicamento*, *sacramento*, *tratamento* ‘to provide – assistance, a course, medicine, a sacrament, treatment’ (Baptista, 1997; Rassi et al., 2016; Rassi, 2023). Naturally, more precise insights on the constructions of this verb ‘*ministrar*’ may be gained through further investigation, potentially by incorporating more annotated data.

Likewise, in the CTA corpus, the verb *cometer* seldom functioned as a *Vsup*, as in (38)

(38) [OP] *Arrepẽdẽdo se do crime que tinha cometido* [mPT] ‘*arrependendo-se do crime que tinha cometido*’ [EN] ‘regreting the crime [he/she] had committed’ [LdM:25447]

This is the exclusive usage of *cometer* in Modern Portuguese, as seen in examples like *cometer* – *crime*, *infração*, *pecado* ‘to commit a crime, infraction, sin’. Its prevailing role in the corpus, however, denotes ‘movement’, as seen in (39), aligning with the Modern Portuguese counterpart *acometer* ‘to attack, approach’:

(39) [OP] *e cõ uirtude do cumo dela uay cometer o basilico e vẽçe o e mata* [mPT] ‘*e com virtude, do cimo dela vai acometer (=atacar) o basilisco e vence-o e mata*’ [EN] ‘“And with valor, from the top of it, he will assail the basilisk, defeating and killing it.”’ [HdE-A:67550]

From a different perspective, since predicative noun can be employed with various support verbs, yielding different meanings, some constructions persist from Medieval Portuguese, while others fell

⁶The ‘#’ separates the *Vopl* from the elementary sentence on which it operates; the indices *i* denote coreference.

in disuse.⁷ For instance, in example (42), *fazer deuacões* ‘fazer devoções’ ‘make devotions’ conveys the notion of offering **prayers**. Conversely, in example (43), *têr deuacão* ‘ter devoção’ ‘have devotion’ expresses the idea of experiencing a **sentiment** towards somebody/something. In Modern Portuguese, both meanings continue to exist.

(42) [OP] *E per todo ho Regno se fazijã muitas deuacões e orações* [mPT] ‘E por todo o reino se faziam muitas devoções e orações’ [EN] ‘And throughout the kingdom, a lot of devotions and prayers were done’ [MISJ:33714]

(43) [OP] *ẽ algũas festas e dias ẽ que tinha mais spicial deuacõ* [mPT] ‘em algumas festas e dias em que tinha mais especial devoção’ [EN] ‘in some feasts and day in which [she/he] had more special devotion’ [MISJ:3650]

In a different context, both *aver doo* in (44) and *tomar doo* in (45) signify having a feeling of pity towards something. In contrast, *fazer doo* in (46) corresponds to an expression of lamentation. In contemporary Portuguese, the latter meaning no longer exists.

(44) [OP] *Doo ey eu da tua morte* [mPT] ‘Dó hei eu da tua morte’ [EN] ‘Pity I have of your death’ [LdM:215464]

(45) [OP] *e tomar doo por ho fallecimẽto desta sancta Senhora* [mPT] ‘e tomar dó pelo falecimento desta santa senhora’ [EN] ‘and take pity for the passing of the holy lady’ [MISJ:45727]

(46) *e disto fezerõ as donas grande doo polla morte de ...* ‘e disto fizeram as donas grande dó pela morte de ...’ ‘and from this, the ladies made a great lamentation for the death of.’ [Vesp:11659]

5 Conclusion and Future Work

This paper systematically explored support-verb constructions (SVC) in the *Corpus de Textos Antigos* (CTA). With a corpus annotated with lemmata and morphosyntactical categories, a little more than

27,000 instances were automatically extracted, involving 81 different candidate verbs, also serving as support verbs. Manual analysis was undertaken to differentiate SVCs from other verb uses, resulting in the identification of a total of 3,094 SVC instances (949 different types). To the best of our understanding, this compilation represents the largest collection of SVCs in Medieval Portuguese currently available. SVCs featuring the verbs *ser* and *estar* (‘be’) are currently under investigation.

From the studied SVC, cases of nominalizations and standalone predicative nouns were identified, along with various support verb variants and types of converse SVC. Distinctive constructions, including operator-verb constructions (both causative and linking operator-verbs), were also identified. Subsequent efforts should aim to systematize the identification of operator-verbs. The most commonly employed support verbs in SVC within this corpus were *fazer* ‘to do/make’ and *haver* ‘to have’. These verbs, being less semantically loaded compared to other support verbs, exhibit a broader distribution concerning the predicative nouns that select them, thereby contributing to their higher frequency of use. Numerous instances of SVC identified in the CTA corpus extend beyond Medieval Portuguese, persisting seamlessly into Modern Portuguese. This sustained continuity solidifies the enduring status of SVC as an integral component of the linguistic heritage within the language. The linear order of SVC constituents remains comparable between both periods, albeit with a few instances in Medieval Portuguese allowing the inversion of Vsup and Npred, a phenomenon that is currently blocked.

Future efforts should prioritize the extraction and analysis of instances (a) based on the list of predicative nouns found at this stage, to explore Vsup variation and discover unseen word senses; and (b) expand it to other predicative nouns determined in Modern Portuguese (Baptista and Mamede, 2020b). Attention should also be directed towards complex noun phrases that arise from the reduction of SVC (Gross, 1981). These phrases are headed by the predicative noun and retain their arguments throughout the reduction process. Furthermore, operator-verb constructions, both causative and linked Vop, related to SVC should be systematically described as well. An examination of the intersection between data related to support verbs and the results obtained for predicative nouns would enhance the diachronic study of Portuguese SVC.

⁷It is also possible to find SVC in which the predicative noun is the *subject* of the sentence (40)-(41) (Baptista, 2022):

(40) [OP] *Ca ẽnos presuntuosos reyna a soberua* [mPT] ‘Porque nos presuntuosos reina a soberba’ [EN] ‘because in the presumptuous reigns pride/arrogance’ [HdE-A:83107]

(41) [OP] *supytamẽte a tomou hũu leue sõpno* [mPT] ‘subitamente, um leve sono tomou-a’ [EN] ‘suddenly, a light sleep overtook her’ [HdE-A:83107]

The Vsup involved in this less studied ‘predicative noun as subject’ construction are the same and support the same type of predicates as found in Modern Portuguese, as the examples above show.

References

- Jorge Baptista. 1997. *Sermão, tarefa e facada: uma classificação das expressões conversas dar-levar*. *Seminários de Linguística 1*, pages 5–37.
- Jorge Baptista. 2005. *Sintaxe dos Predicados Nominais com ser de*. Fundação para a Ciência e a Tecnologia & Fundação Calouste Gulbenkian, Lisboa.
- Jorge Baptista. 2022. Support verb constructions with predicate noun in subject position. *Bulletin de linguistique appliquée et générale*, 40:379–397.
- Jorge Baptista and Nuno Mamede. 2020a. *Dicionário Gramatical de Verbos do Português Europeu*. Universidade do Algarve.
- Jorge Baptista and Nuno Mamede. 2020b. Syntactic Transformations in Rule-Based Parsing of Support Verb Constructions: Examples from European Portuguese. In *9th Symposium on Languages, Applications and Technologies (SLATE 2020)*, volume 83 of *OpenAccess Series in Informatics (OASIs)*, pages 11:1–11:14, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.
- Maria Inês Bico, Jorge Baptista, Fernando Batista, and Esperança Cardeira. 2022. Early experiments on automatic annotation of portuguese medieval texts. In *Linking Theory and Practice of Digital Libraries*, pages 442–449, Cham. Springer International Publishing.
- Ivo Castro. 2006. *Introdução à História do Português*, 2.^a edição. Edições Colibri.
- Lucília Chacoto. 2005. *O Verbo Fazer em Construções Nominais Predicativas*. Ph.D. thesis, Universidade do Algarve, Faro.
- Lucília Chacoto. 1996. Predicados nominais com *fazer* no português medieval. In *Actas do XII Encontro da Associação Portuguesa de Linguística*, volume 2, pages 69–77.
- Aggeliki Fotopoulou, Eric Laporte, and Takuya Nakamura. 2021. Where do aspectual variants of light verb constructions belong? In *Proceedings of the 17th Workshop on Multiword Expressions (MWE 2021)*, pages 2–12. Association for Computational Linguistics.
- Gaston Gross. 1989. *Les construction converses du français*. Droz, Genève.
- Gaston Gross. 2022. *Manuel d’analyse linguistique: approche sémantico-syntaxique du lexique*. Presses universitaires du Septentrion.
- Maurice Gross. 1981. Les bases empiriques de la notion de prédicat sémantique. *Langages*, 15(63):7–52.
- Maurice Gross. 1996. Lexicon-grammar. In Keith Brown and Jim Miller, editors, *Concise Encyclopedia of Syntactic Theories*, pages 244–259. Pergamon, Cambridge.
- Zellig Sabettai Harris. 1964. The elementary transformations. In Henry Hiz, editor, *Papers on Syntax*, pages 211–235. D. Reidel Pub. Co.
- Zellig Sabettai Harris. 1991. *A Theory of Language and Information. A Mathematical Approach*. Clarendon Press, Oxford.
- Maarten Janssen. 2016. [TEITOK: Text-faithful annotated corpora](#). In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC’16)*, pages 4037–4043, Portorož, Slovenia. European Language Resources Association (ELRA).
- Maria Regina Pante. 2012. O verbo *tomar* como verbo-suporte no português arcaico. *Línguas & Letras*, 13(24).
- Maria Regina Pante and Érica Fernanda Ortega. 2009. O verbo *tomar* como verbo-suporte no livro de ensinância de bem cavalgar toda a sela. *Revista Philologus*, 125:144–152.
- Elisabete Ranchhod. 1990. *Sintaxe dos predicados nominais com estar*. Instituto Nacional de Investigação Científica (INIC).
- Elisabete Ranchhod. 1999. Construções com nomes predicativos na *Crónica Geral de Espanha de 1344*. In I. H. Faria, editor, *Homenagem ao Homem, ao Mestre e ao Cidadão*, pages 667–682. Edições Cosmo.
- Amanda Rassi, Nathalia Calcia, Oto Araújo Vale, and Jorge Baptista. 2016. Estudo contrastivo sobre construções conversas em PB e PE. In O. Nadin, A. Ferreira, and C. Fargetti, editors, *Léxico e suas interfaces: descrição, reflexão e ensino*, pages 199–218. Cultura Acadêmica.
- Amanda Pontes Rassi. 2023. *O verbo dar em português brasileiro. Descrição, classificação e processamento automático*. Letraria, São Paulo.
- Helmut Schmid. 1994. Probabilistic part-of-speech tagging using decision trees. In *Proceedings of International Conference on New Methods in Language Processing*, pages 154–163.
- Helmut Schmid. 1999. *Improvements in Part-of-Speech Tagging with an Application to German*, pages 13–25. Springer Netherlands, Dordrecht.
- Gael Vaamonde and Maarten Janssen. 2020. *Da edición dixital á análise lingüística. A creación de corpus históricos na plataforma TEITOK*, pages 271–292.
- Rita Veloso. 2013. Gramática do português. volume 2, chapter Subordinação relativa, pages 2061–2134. Fundação Calouste Gulbenkian.

Appendix A - SVC from the CTA corpus

ID	Right context	Vsup	Left Context
MISJ:34147	afeytauas afeytou sse ela . Quando tu	aauias	fame ouue ela fame e quando tu
MISJ:38305	que nos chamamos filho de deus .	auêdo	mjsericordia sobre os homêes
HdE-A:117731	muitas lides , e contendas as de	auer	com o inimigo , ca sei sarta
LdM:117037	que os que pecã nõ	auerem	pena . E estes maaes sobreditos
LdM:38822	e preguar , pera as gentes	auerem	notiçia , e conhocimento a sua vida
MISJ:13404	o mestre salla preguntou a jacob onde	aueria	pousada e jacob respondeu lhe e dixe
LdM:124182	E porque algũ per vêtura	aueria	sabor de os aprêder teue por bẽ
LdM:251724	mas empero, que grande fuza ella	auia	em Deos , nenhũ nõ sabe pero
LdM:147565	comgrande alegria porque	auia	reçeo do padre encorrer na sua ira
LdM:189484	pera fazerem festa , assi como	auiam	custume de fazer quada sabodo
LdM:66406	sobre esto cõ hũũ sancto bispo que	auja	nome alexandre
MISJ:14442	Mais porque o sancto bispo nõ	auja	aquella arte nem tal sçiencia
MISJ:25178	Exemplo Hũũ homẽ	auja	grande odio mortal a outro
MISJ:9536	assy . que este homẽ que asy	auja	este malquerença
LdM:133001	contra ele com grande sanha que	auya	cuydãdo em que gujsa
VM-G1:4363	Ayda muito minyno pera Casar e	aver	erdeyro e ser sempre doente
MISJ:7006	neste moesteiro . pera que	averia	êteyro poder
LdM:18116	papa e mestre da ordẽ .	Avido	maduro cõsselho . e vêedo seer
HdE-A:91771	lenco na mão e Rostro . nõ	cessãdo	de lagrimas e salucos e vrrros mũy
Vesp:338	suplicou ao sancto padre lhe	cõcedesse	graca de plenaria yndulgẽcia
LdM:104132	aquelle era seu filho . E mãdou	cometer	casamêto cõ elle pera sua filha
LdM:191166	Arrepêdêdo se do crime que tinha	cometido	bolueo as trinta peças de prata
LdM:82187	soo filha do dito Conde . lhe	Cometyã	grãdes e altos Casamêtos cõ duques
VM-G1:492	pera lhe seer per elles	Concedido	e outorgada licêca e liberdade pera
Vesp:3975	daquella terra treeuosa e	cuberta	d' escuridade e de morte perdurauel
LdM:197570	E porque leyxou a sabedorya	ẽcorreo	ẽ ygnorancia . que he abetamento
LdM:27972	sustancia . nõ departida .	emcorreo	ẽ cobijça que he corrupcõ que
HdE-A:30929	da mête e da alma e faz	falecer	a ujsta do verdadeyro conhecimêto
HdE-A:122634	que aquesto fosse . aynda aly nõ	faleçia	a ajuda . nem a merçe de
LdM:74342	os homẽs . E com aquelle ujnho	fara	o senhor deus cõujte ao seu pouoo
MISJ:18828	grãde tempo que o nõ vio .	faz	grãde alegria . Assy se alegrauã eles
LdM:106925	esto faz aquelle que nõ cuyra nẽ	faz	conta que os outros cuyrem delle .
LdM:40977	todos aqueles que	fazẽ	a sua uõtade . Elle destruy cruelmête
VM-G1:10999	sã çebriã começou a chorar	fazêdo	cõfissã de todolos peccados
LdM:17088	se ally gastã aquelle dia . E	fazem	muytas outras alegrias
HdE-A:145171	e seus amjgos . tomã prazer e	fazem	grandes convites . E esto ham elles
MISJ:30459	todos a fogir asanhãdo se . e	fazendo	muy grãde arroydo . e disserom todos
HdE-A:107979	o bispo de yterãna	fazendo	bõa vida mereçeo de se gozar as
MISJ:8335	gente . chorando e	fazendo	chanto a muy grandes vozes . por
HdE-A:94185	aa sua madre muy sancta . e	fazendo	conparaçõ antre ty e elles acharas que
Vesp:2264	toda ajuda que fisicos lhe podiã	fazer	E metera sse a handar per
MISJ:39328	Outrossy nõ deue homẽ deixar de	fazer	a caridade aos Jrmããos pello studo da
MISJ:6126	seruiã outrossy a ella . sẽ	fazer	cerymonyas nẽ deferêca
HdE-A:58119	a castigar . e afaagar . e	fazer	lhe muytos viços . e a falar
HdE-A:120357	ẽ aquela festa começou de	fazer	tã grãdes virtudes e tã grãdes marauilhas
VM-G1:3723	a qual disse . molher boas obras	fazes	: leuãta te pella manhaã e sairas
LdM:137425	ante que fosse abade	fazia	grandes abstêêças
VM-G1:13442	homẽ bebedor . e que	fazia	adulterio . e amigo de totalas cousas
Vesp:8701	he casado cõ ela aquele que lhe	fazia	muytos falsos afagamêtos primeyro
LdM:211701	que lhe derõ pollo bẽ que lhes	fazia	Assy senhor que eu nõ sey
MISJ:1210	os tormentos que sofriam . Mas pero	fazia	lhes deos muyto bem e muyta merçee
HdE-A:54198	assi todos na igreja pella chuiua que	fazia	, hũa bõa molher auendo doo de
HdE-A:48044	atee os mesmos caualleiros	faziã	bulrra dele . e se chegauã a
MISJ:24022	de sam frascisco . os frades	faziam	ujda estreyta e aspera ẽ grande pobreza
VM-G1:5367	a el rrey d'ugaria . e foy	fecto	o casamêto da madre cõ o filho
MISJ:1682	que vio o grãde aluoroço que era	feito	ẽno poboo mandou o de cabo emçarrar
MISJ:7766	E em aquele logar som muytos beês	feitos	por ele a louuor de nosso senhor
HdE-A:68016	Estas tres batalhas forã	feytas	ẽ tres ãnos . en que forã
HdE-A:92056	Como se solêpnemête teuesse	feyto	voto de proffysam e obediencia
MISJ:45134	Ja entõ nõ podendo mais	fez	ho padre a absoluycã
HdE-A:61653	de çiriaco a carpasio o vigairo e	fez	banho ẽna fonte do bauptismo
HdE-A:31394	das brauezas que o adiantado	fez	contra o seu corpo . e todolos
HdE-A:150910	aquelle que a muytos parêtes se nõ	fezer	sua uõtade e seu prazer delles .
LdM:143707	totalas chagas que [. . .] maximiano	fezera	em my . e esto com o
LdM:164715	e door . screueo lhe que nõ	fezesse	aballo nẽ partisse . atee elle lho
LdM:164095	pallauras . e discretas .	ffez	hũa aRenga ante el Rey . princepe

Appendix B- List of predicative nouns (different types) in SVC

aazo 2, aazos 1, abaixamêto 1, aballo 4, abastâca 1, abastança 1, absolucyã 1, absteêças 1, abstinências 1, accjentes 1, acordo 4, adulterio 2, adulterios 1, afagamêtos 1, afeycã spiritual 1, afeycom 1, affeycam 1, agrauro 1, âgustia 1, ajuda 4, ajuntamento 1, aJütamêto 2, alcouuetaria 1, alegria 11, alegrias 2, allegria 1, almoeda 1, aluoroço 2, amargura 1, ameaças 2, amizade 1, amor 10, andãca 1, andanca 2, andança 4, angustia 1, aparelho de oracô 1, aRenga 1, armas 1, armas de caualaria 1, aroido 2, arroido 2, arroydo 8, arte 2, artes 1, asesejo 1, asessejo 1, assêto 1, astença 1, astençaa 1, auareza 1, auerya 1, auêturança 1, auito 1, auôdamêto 1, auôdança 3, auôôdança 1, auorrecimêto 1, auto da caualaria 1, autos 1, avito 4, avito da Religiã 1, avito da Relligiam 1, avyto 4, avyto da Religiã 1, avyto da Relligiã 1, avyto de Religiã 2, avyto de Relligyã 1, banho 5, batalha 1, batalhas 2, bathalha spiritual 1, bẽ 4, bẽcam 1, beens 1, beês 5, bẽs 3, bem 23, bem auêturãca 3, bẽs 1, beuedice 1, blandeza 2, bõa vida 1, boas manhas 1, bõas manhas 1, boas obras 1, bõns costumes 1, booa obra 1, boas obras 1, bõs costumes 1, brauezas 1, brauura 1, brazer 1, bulra 1, bulrra 1, buracos 1, cabo 1, calma 1, caminho 1, caridade 4, carraduras 1, Carrego 1, carta de reposta 1, Cartas de alfforrya 1, casamêto 3, Casamêtos 1, castidade 4, castydade 1, caualaria 2, cerco 1, cerymonyas 3, chaeza 1, chaga 3, chagas 5, chamas 1, chanto 1, choro 1, choros 1, chuiua 4, ciêcias seglaes 1, cienci 1, cnydado 1, coacão 1, cobijca 1, cobijca 3, côcelho geeral 1, cõfissã 1, cõfusom 2, cõfyssam 1, começo 1, conselheiro 1, comselho 2, comtenda 1, concelho (=conselho) 1, concelho géral (=conselho geral) 1, confeyções de blandeza 1, confiança 1, confissao 1, Confusã 1, Confyãca 1, conhecimêto 4, conhocimêto 1, conpaixom 1, companhia 1, comparacô 2, comparacô 1, comparacom 1, comparaçom 1, conselho 11, consolações 1, consollações 1, Conselheiro 2, conta 6, contêda 1, contenda 3, contendas 2, contreadições 1, convites 1, côroes 1, cõpaixô 1, coraçã 1, coracô 1, coraçô 2, coraçom 3, coracôões 1, coroa 1, cortesias 2, cõselho 19, cõsolacô 1, cõsolacão 3, cõsollacô 1, cõsollacôe 1, cõsollacôões 1, cõselho 1, cõsollacô 2, costume 5, cõta 1, cõtêda 2, cõtentamêto 1, cõteteiro 1, cõujte 2, cousa de vytuperyo 1, coydado 1, coyta 4, creença 2, cremça 1, crime 1, crueldade 5, crueza 1, cruezas 1, cruezes 1, çugidade da luxuria 1, cuidado 26, cuidoado 2, culpa 4, culpas 1, cura 3, cura de regno 1, curação 1, custume 8, costumes 6, cuydações 1, cuydado 8, cuydados 1, cuyta 1, Cyrymonjas 1, dano 5, dano 5, danos 2, dapno 2, dâpno 1, deciplinas 1, deferêca 2, degredo 2, deleitaçô 1, deleitaçom 1, deleitaçom terreal 1, deleitações 2, deleitações 1, deleitos 1, deleytaçom da carne 1, deleytações 1, deleytos 1, demostramêto 1, demudamêto 1, departimêto 3, descãso 1, desde 2, desdem 2, desefadadêto 1, desejo 1, desenfadadêto 1, desfalecimêtos 1, desonrra 5, desonrras 1, desõrra 2, desõrras 1, despeito 1, despesas 3, despessa 1, despessas 2, desprazer 2, disputaçom 1, destruyamêto 1, desuayro 1, detêça 1, detreminacã 1, deuacã 1, deuacô 1, deuacô 1, deuacões 2, deuer 1, dia 1, dignidade 5, dignidades 1, dignjdade 6, dignjdades 1, dileitações corporaes 1, dileytações 1, diligência 1, dilligência 1, direyto 1, disciplina 1, discordia 1, dom 1, doo 26, dóo 3, door 5, dóor 2, door jnyrya scarnho 1, dor 1, dores 2, doutrina 1, duida 1, dulcura 2, dulçura 1, duuida 1, duuija 1, êbargo 4, êbargos 1, êbates 1, êfermidades 1, êfirmiyade 1, êJuria 4, êJurias 4, emcantamêtos 1, emêda 1, emêda 2, enfermydade 3, emfirmjidade 3, emjizade 1, emJuria 3, emmenda 3, emsinamêtos 1, emtêdimêto 1, emtendimento 1, emtendimêto 1, entrada 1, enmueja 2, enveia 2, enfadamento 1, enfirmjidade 1, engano 1, engenho 2, enJurias 1, entêdimêto 1, entêdymêto 1, entêçom 1, entrada 1, enueja 1, enveja 1, êpresa 2, erdeyro 1, erro 3, error 1, escanhos 1, escarneio 3, escarneos 1, escarnho 8, escarnhos 2, escarnio 2, escuridade de morte 1, escuridade 1, esmolas 4, esmolla 2, esmollas 2, espanto 6, espãto 3, esperãca 5, esperança 9, experiencia 1, estimacão 1, estoria 1, estrago 1, estudo 1, êtêdimêto 2, êtêçom 1, êtendimêto 3, êueya 2, execuçô 1, exemplos 1, exenplos 1, exêplo 1, facanhas de doo 1, falamêtos 1, falla 1, falla (=fala) 1, fama 3, fame 10, famjliaridade 1, fastidio 1, fe 5, febre 2, fecto d'armas 1, fedor 1, fee 4, fehuza 2, feira 1, feitos 1, fel 1, ferydas 1, festa 23, festas 4, feuz 1, feyto 1, ffe 1, ffym 1, ffym (fim) 1, fiêdade 1, fihuza 2, filhos 1, fim 6, firmeza 2, fiuza 13, fogo 8, fogos 1, fojos 1, folganca 2, folganca 2, força 6, forca (=força) 2, forcas corporaes 1, forro 1, forteleza 1, fremosura 2, fremosura 3, fremosura spiritual 1, frio 3, friu 1, fruito 2, frujto 1, fruyto 5, fruyto de boas obras 1, fruyto de mansidoem de odor 1, fugeira 1, fumo 1, furaco (=buraco) 1, furto 2, furto (fruto) 1, fym 1, fyuza 2, galardom 1, geada 2, géeracom 1, geesto 1, gestos de cabeça 1, gloria 5, gosto 1, gracia 5, graça 2, guerra 10, guerra de fogo 1, guerras 1, herdeyro 2, hidade 1, ho que 1, homicidio 1, honra 2, honrra 4, honrras 1, hõra 1, hõras 1, horo (=oração) 1, hõrra 25, humor 1, infirmydade 1, iniura 1, injuria 1, intelligência 1, iogos 1, iustica 1, Jantar 1, Jeyto/door: Jeyto algũu de door 1, Jeyto/door: Jeyto nêhiu de door 1, Jinchacos postemas 1, Jnfamja 1, Jnvirias 1, jogos 3, juizo 1, Jujzo 1, juramêto 1, Justica 1, Justica 7, Jütamêto 1, lagrimas 1, lanbugem 1, lecença 1, letera 1, letras 3, ley 1, leytações 1, leytura 1, liberdade 1, licêca 3, liçêca 1, licêca spiciall 1, licom 1, lides 1, linhagem 1, loucaynha 1, loucura 1, louour 3, lououres 1, lume 2, luxuria 2, mãã vida (mã vida) 1, maaes (=males) 1, maa rosto 1, maa rostro 1, maaos feytos 1, mãas cousas 1, maas obras 1, mãdado 2, mãdamêto 1, mal 119, maldade 4, maleficio 1, maleficios 2, males 15, malicia 1, malquerença 1, mandado 2, mandamento 2, mandamêto 1, manyha 1, marauilha 2, marauilhas 22, marauilhas 2, marteiro 6, marteiro 2, mãyar de fogo 1, mãyar do uermẽ 1, mazella 1, medo 72, meester 1, meestrado 1, meezinha 1, meezinhas 1,

memoria 1, menagẽ 3, mencô (=menção) 1, menêcoria 1, menemquoria 1, mentes 5, merçe 13, merçed 1, mercee 5, merçee 24, mercees 2, mercecimêto 1, mercecimêtos 1, meserecordia 1, mester 64, mester sabeloria 2, mête 1, mêtês 7, mezuindade 1, migoa 1, migua 1, milagre 6, milagres 12, milagros 1, millagre 22, millagres 21, mingoa 2, mingua 2, misericordia 3, missa 1, mjoas 1, mjgua 2, mjlagres 1, mjsericordia 2, molher 1, morada 3, morte 3, mortes 1, mostrancas 1, moujmêtos 1, mudaçô 1, mudamêto 1, mudanca 1, mgoa 2, natureza 1, necessidade 1, neçessidade 1, njnho 3, nobreza 3, nojo 5, nome 400, nomẽ 2, nomo 1, noticia 2, notiçia 1, noyos (=nojos) 1, noyte 1, obediçya 1, obediencia 1, obra 11, obra de geeraçom 1, obras 7, obras de bem 1, obras de piedade 1, occupacô 1, odi 1, odio 7, odio mortal 1, odor 4, offêças 1, officio 6, offiçio 1, officios 2, officios bayxos 1, oficio 3, ofiçio 1, oportunidade 2, oraçã 14, oraçam 6, oracão 13, oraçã 5, oraçã 12, oracô 2, oraçô 38, orações 6, orações 5, oracom 1, oraçom 36, oraçon 1, orações 1, oramêto 1, oras 1, orões (=orações) 1, paciência 3, paciencia 2, paixã 2, paixô 1, paixom 1, paixôes 1, palaura 1, palavras 1, pallauras 1, parauoas 1, parte 7, partimêto 1, pascoa 1, patimêto 1, pawor 2, paz 2, pccado da luxuria 1, pccado 3, pccados 2, pccado 9, pccados 5, peêdêca 2, peêdêca 2, peêdença 2, peêdença 6, peendença 2, pena 4, pênã 2, penas 4, pendêca 1, pendêca 2, pendença 1, penitência 3, penitência 1, penitencia 3, penitencia 1, perdimêto 1, perdô 1, perseguiçã 1, perseguiçô 2, perseguiçôes 1, perseguiçôes 1, persiguiçã 1, persiguiçô 4, persiguiçôes 2, persiguiçom 1, persyguicô 1, pêsamêto 1, pesar 25, pescaryas 1, petycam 1, piedade 10, pinitência 1, planto 7, plãto 3, placer 2, pobreza 3, podeiro 1, poder 26, poderio 17, poderios 2, poderyos 1, portas 1, pousada 1, pranto 1, pratica 1, pratyca 2, prazer 48, prazer fingidico 1, prazer spiritual 2, prazeres 2, prea 1, preegaçã 1, preegaçô 1, preegaçô 1, preegaçô 1, preegaçô 1, preegaçô 1, preito 1, preposyto 1, pressa 1, preytasia 1, prezes 1, procyssã 1, procyssões 1, proffysões 1, proffysam 1, proffysã 3, proffysam 1, prologo 1, promessas 1, prometimêto 3, proposyto 1, proueito 9, proueyto 5, proueytos 1, proveyto 1, quarentura 1, quareemas 1, quebrãto 1, quêçtura 1, queymas 1, queyxume 1, quietacã 1, raiz 1, rayua 1, rayzes 1, razô 2, razões 2, receo 1, reço 1, receyo 1, relãpagos 1, remedios 4, renêbranca 1, Repartimêto 1, repêdimêto 1, reposta 4, repostas 1, Repouso 3, reuerêca 2, reuerêca 2, Reuerêca obediência 1, Reuerêcas 1, reuerencia 3, reuerença 1, reuerencia 1, reuerencia 1, rezões 1, rigor 1, rogo 1, rogos 3, roido 1, roubo 1, royo 3, rreposta 1, rrigor 1, saão 1, sabeloria 8, sabelorya 1, sabendas 1, saber 2, sabor 18, sacrefiçio 12, sacrefiçio 2, sacrefiços 2, sacrificio 14, sacrificio 45, sacriçio 2, sacriçios 5, sacriçios 13, sacriçios 1, saluaçom 1, salucos 1, samha 1, sandices 1, sangue 1, sanha 43, santidaade 1, saude 10, saude da alma 1, sciência 2, sciencia 3, scusações 1, sebe 1, segura 1, sede 7, seguraça 1, segurança 2, semelhança 1, senhorio 5, sepultura 2, serãao 1, seraões 1, sermô 4, sermom 1, seruêtiya 1, seruico 2, seruico 12, seruicos 2, seruicos 1, seruissô 1, seruico 2, servissô 1, sesejo 1, sêtimêto 1, signaaes 1, signal 1, signal da cruz 1, sinaaes 1, sinaes 5, sinães 1, sinal 3, sinal da crus 1, sinal da cruz 8, sinal de acenamêto por falla 1, sinhaes 2, sinhal 1, siso 1, soberua 2, soffrença 1, sofrêça 1, soffrença 1, solaz 1, solazes 1, sollaz 1, sollazes 1, sollêpnydade 2, sôo (=som) 1, sôpno 6, sorte 2, sospeita 1, sospeyta 1, sotilezas 2, spanto 1, spãto 2, sperãca 1, speranca 3, sperança 4, ssôo (som) 1, ssynal 3, ssynal da Cruz 1, estabilidade 1, stado 1, stimaçom 1, studo 1, suplicaçom 1, suspeyta 1, sylêcio 1, sylêcio 1, synaes 1, synal 3, synal da cruz 9, synhal 1, talante 2, tallante 3, tauoada 1, teebras 1, temor 30, temores 1, tempo 6, tempos 2, tenperança 1, têpestade 1, têpo 3, têpos 4, termo 1, termos 1, testamento 1, testamêto 2, testemunho 1, testymunho 1, têtacô 1, themor 2, titulo 1, tormento 1, tormentos 6, tormêto 2, tormêtos 5, torto 12, tortos 5, toruacô 1, toruacô 1, toruom 1, trabalho 8, trabalhos 7, trahycões 1, trayçô 1, trayções 1, traycom 1, trebelhos 2, trebuto 1, treyçam 1, treyçom 1, tribulações 1, tribulações 1, tribulações 1, tristeza 4, tristura 1, tromentos 1, tryuimêto 1, uããgloria 1, uaydades 1, uentuyra 1, uergonca 1, uergonça 10, uergonha 1, uersos 1, Uersus (=versos) 2, uêtura 2, uida 5, uigãca 1, uirtude 5, uijos 2, uijos de comeres 1, ujda 1, ujança 1, ujsta 2, uodas 4, uodas (=bodas) 1, uôôdade 3, uoto 1, uso 1, uyngança 1, vagar 1, vaydade 1, vegalias 1, vergôca 4, vergomça 2, vergonça 5, vergonha 1, versos 1, virtudes 4, viço 1, vicos 1, viços 2, victorias 1, victorya 2, vida 23, vida de Religiã 2, vigalias 1, vigillias 1, vingãca 6, vingança 1, vingança 1, virtude 1, uirtude de auôdança 1, virtudes 9, vista 1, visytacões 1, vitoria 3, vjda 7, vodas 1, vontade 3, vontades 1, voontade 2, voôtade 7, voôtades 1, vôtade 3, vôtades 1, voto 5, voto de proffysam 1, votos 3, vrrros 1, vrsã 1, ffinta 1, ygnorancia 1